



**Gonalo Maia Jcome da Costa**

Licenciado em Cincias de Engenharia  
Electrotcnica e de Computadores

**Map-it –  
Mapeamento Tridimensional de Interiores**

Dissertao para obteno do Grau de Mestre em  
Engenharia Electrotcnica e de Computadores

Orientador: Tiago Cardoso, Professor Doutor, FCT-UNL

Jri:

Presidente: Professor Doutor Rui Manuel  
Leito Santos Tavares – FCT/UNL

Arguentes: Professor Doutor Andr Teixeira  
Bento Damas Mora – FCT/ UNL

Vogais: Professor Doutor Tiago  
Oliveira Machado de  
Figueiredo Cardoso – FCT/UNL



FACULDADE DE  
CINCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA

**Setembro, 2014**

## **Map-it**

Copyright © Gonalo Maia Jcome da Costa, Faculdade de Cincias e Tecnologia, Universidade Nova de Lisboa.

A Faculdade de Cincias e Tecnologia e a Universidade Nova de Lisboa tm o direito, perptuo e sem limites geogrficos, de arquivar e publicar esta dissertao atravs de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar atravs de repositrios cientficos e de admitir a sua cpia e distribuio com objectivos educacionais ou de investigao, no comerciais, desde que seja dado crdito ao autor e editor.

*A principal meta da educação é criar homens que sejam capazes de fazer coisas novas, não simplesmente repetir o que outras gerações já fizeram. Homens que sejam criadores, inventores, descobridores. A segunda meta da educação é formar mentes que estejam em condições de criticar, verificar e não aceitar tudo que a elas se propõe.*

- Jean Piaget

... a todos quantos gostariam de ter a oportunidade de poder aprender e a todos os que, como pais e professores, através do seu exemplo pessoal são facilitadores da construção pessoal de quem se encontra em crescimento...

Gonçalo Maia Jácome da Costa



# Agradecimentos

Gostaria de agradecer ao meu orientador, Prof. Tiago Cardoso, por ter acreditado no meu projecto e me ter dado a oportunidade de o desenvolver, assim como por todo o apoio e motivação ao longo desta fase de preparação para a dissertação que me foi tão importante.

À Universidade Nova de Lisboa e à Faculdade de Ciências de Tecnologias em particular, por me terem acolhido e disponibilizado a conquista dos recursos que me serão úteis na minha vida profissional e pessoal, dado que foi aqui, com o apoio do corpo docente, dos meus professores directos e dos múltiplos colegas que cruzaram o meu caminho, particularmente aqueles que se tornaram amigos e grandes companheiros nesta caminhada, que cresci de uma forma muito especial. A todos eles de uma forma geral, o meu agradecimento sincero.

À minha irmã e cunhado agradeço o exemplo de perseverança e luta para conquistar objectivos de vida. Aos meus pais, um agradecimento muito particular, por me terem proporcionado a possibilidade deste percurso académico e por todo o apoio, ânimo e conselhos esforçados, sobretudo nos momentos mais difíceis. Aos restantes familiares pelo carinho que sempre me deram.

Obrigado.

Gonçalo Maia J. C.

Sexta 20 de Setembro, 2014



# Abstract

Three-dimensional perception is a growing field of study. Although there has been emerging studies about the application of different kinds of technologies to obtain models that fit to the reality, a technology with an ideal perception to all kinds of scenarios does not exist yet.

In this dissertation the development of a portable low-cost three-dimensional system was proposed, with a wide range of applicability, so it can perceive low light indoor spaces, such as building rooms or underground mines.

The three-dimensional data from a scenario was acquired using triangulation between a digital camera, laser line points and a laser beam. The use of digital image processing was a resource required to build the sceneries models.

This dissertation validation consisted of the realization of different tests of the prototype so it would be determined its perceptual abilities. Firstly a set of calibration tests where specific points of the scenario were studied individually multiple times in order to evaluate the sensor's precision. Afterwards it was studied the ability to transform the projected laser line into a group of coordinates. Finally, the ability to three-dimensionally perceive the scenario was tested by rotating the sensor to make the transformation of multiple line projections into a group of coordinates around the sensor.

Results show that on the calibration test the acquisition of distances had a relative mean error below 1%, that on the laser line test it had the ability to perceive the depth of multiple points on that line and on the rotation test it was possible to acquire depth lines sets from the scenery by rotating the sensor.

The results obtained in the validation tests allowed to conclude the viability of using triangulation for the acquisition of coordinates sets on indoor spaces.

KEYWORDS - THREE-DIMENSIONAL PERCEPTION, INDOOR MAPPING, LOW-COST, PORTABLE, 3D SCANNER, TRIANGULATION



# Resumo

A percepção tridimensional é uma área em crescente desenvolvimento. Não existindo uma tecnologia ideal capaz de perceber todo o tipo de cenários, têm surgido estudos sobre a aplicabilidade de diferentes tecnologias de forma a obter modelos que melhor se aproximem da realidade.

Propõe-se nesta dissertação o desenvolvimento de um sistema de baixo-custo de percepção tridimensional que seja portátil e acessível na sua aplicabilidade, de forma a ter a capacidade de perceber tridimensionalmente espaços interiores com pouca iluminação, tais como as divisões de um edifício ou os canais de uma mina subterrânea.

Utilizou-se a triangulação entre uma câmara digital, os pontos de uma linha laser e um emissor laser para a obtenção dos dados tridimensionais do cenário. Para isso, recorreu-se a metodologias de processamento digital de imagens para a construção dos modelos dos cenários.

Para se obter a validação desta dissertação, realizaram-se testes do protótipo de forma a determinar as suas capacidades percecionais. Em primeiro lugar realizou-se uma bateria de testes de calibração onde, repetidamente, se estudaram isoladamente pontos do cenário para verificar a precisão do sensor. Após esta calibração, estudou-se a capacidade de transformação de uma linha laser projetada num conjunto de coordenadas. Finalmente testou-se a capacidade de percepção tridimensional por rotação do sensor, de forma a realizar a transformação de múltiplas projeções de linhas laser num conjunto de coordenadas em torno do sensor.

Verificou-se: no teste de calibração, a obtenção de distâncias com erro médio relativo inferior a 1%; no teste da linha laser, a capacidade de percepção de profundidade de múltiplos pontos ao longo da mesma; no teste de rotação, a aquisição de múltiplos conjuntos de linhas de profundidade no cenário, por rotação do sensor.

Os resultados dos testes de validação permitiram concluir a viabilidade de utilizar a triangulação na aquisição de conjuntos de coordenadas de espaços interiores.

PALAVRAS-CHAVE - PERCEÇÃO TRIDIMENCIONAL, MAPEAMENTO DE INTERIORES, BAIXO-CUSTO, PORTÁTIL, SCANNER 3D, TRIÂNGULAÇÃO



# Índice

<b>Dedicatória</b>	<b>ii</b>
<b>Agradecimentos</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Resumo</b>	<b>vi</b>
<b>Índice</b>	<b>ix</b>
Lista de Figuras . . . . .	xii
Lista de Tabelas . . . . .	xiii
Lista de Símbolos e Acrónimos . . . . .	xv
<b>1 Introdução</b>	<b>1</b>
1.1 Enquadramento . . . . .	1
1.2 Motivação . . . . .	2
1.3 Objectivos . . . . .	3
1.4 Sinopse . . . . .	3
<b>2 Estado da Arte</b>	<b>5</b>
2.1 Métodos de Aquisição . . . . .	5
2.1.1 Contacto . . . . .	5
2.1.2 Não-Contacto . . . . .	7
2.1.2.1 Passivo . . . . .	7
2.1.2.2 Activo . . . . .	8
2.2 Processamento de Imagem . . . . .	13
2.2.1 Ruído em Imagens Digitais . . . . .	14
2.2.2 Filtros de Remoção de Ruído . . . . .	15
2.2.2.1 Filtros Lineares . . . . .	17
2.2.2.2 Filtros não-Lineares . . . . .	18
2.3 Segmentação . . . . .	19
2.4 Sinopse . . . . .	21
<b>3 Proposta</b>	<b>23</b>
3.1 Processo . . . . .	23
3.2 Hardware . . . . .	26

3.3	Software . . . . .	28
3.3.1	Sistema de Comunicação . . . . .	28
3.3.2	Sistema de Visão Computacional . . . . .	30
3.3.2.1	Aplicação de Filtro . . . . .	31
3.3.2.2	Passagem para Escala de Cinzentos . . . . .	31
3.3.2.3	Segmentação . . . . .	32
3.3.2.4	Seleccção da Área do Laser . . . . .	33
3.3.2.5	<i>Laser Pinpoint</i> . . . . .	34
3.3.2.6	Reconstrução do Laser . . . . .	36
3.3.3	Sistema de Cálculo . . . . .	40
3.4	Sinopse . . . . .	40
<b>4</b>	<b>Validação</b>	<b>41</b>
4.1	Método de Validação . . . . .	41
4.2	Protótipo . . . . .	42
4.2.1	Cálculo das distâncias . . . . .	44
4.2.2	Cálculo do Ângulo da Câmara . . . . .	47
4.3	Experimentação . . . . .	49
4.3.1	Teste de Calibração . . . . .	50
4.3.2	Teste de percepção de profundidade de uma linha laser . . . . .	51
4.3.3	Teste de percepção tridimensional por rotação do sensor . . . . .	53
4.4	Sinopse . . . . .	54
<b>5</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>55</b>
5.1	Conclusão . . . . .	55
5.2	Trabalhos Futuros . . . . .	56
	<b>Bibliografia</b>	<b>57</b>



# Lista de Figuras

1.1	Actual cobertura da aplicação " <i>Google Street View</i> " no mundo. . . . .	2
2.1	Exemplo de um sistema CMM. . . . .	6
2.2	Esquema de funcionamento de um <i>scanner</i> 3D utilizando triangulação. . . . .	9
2.3	Alguns exemplos de padrões luminosos utilizados . . . . .	9
2.4	Esquema de relação entre a profundidade relativa e a disparidade medida. . . . .	10
2.5	Exemplo de dados utilizados para determinação de distâncias usando <i>Time-of-Flight</i> (Tempo-de-Voo) (ToF). . . . .	11
2.6	Representação dos métodos de aquisição de distâncias ToF(a) e <i>Phase Shifting</i> (Comparação de Fase) (PS)(b). . . . .	12
2.7	Exemplo de dados utilizados para determinação de distâncias usando PS. . . . .	13
2.8	Comparação entre duas imagens com e sem ruído Gaussiano. . . . .	14
2.9	Imagem afectada pelo ruído <i>Salt-n-pepper</i> . . . . .	15
2.10	Mecânica de funcionamento de filtragem espacial usando uma máscara de tamanho 3x3. . . . .	16
2.11	Exemplo de aplicação de um filtro de média linear. . . . .	17
2.12	Exemplo de aplicação de um filtro de mediana com uma máscara 3x3. . . . .	18
2.13	Exemplo de análise de histograma para determinação de valor de <i>threshold</i> . . . . .	19
2.14	Exemplo de utilização de <i>edge detection</i> . . . . .	21
3.1	Esquema geral do funcionamento do processo proposto. . . . .	25
3.2	<i>Concept</i> do sensor a utilizar no processo. . . . .	27
3.3	Acesso através de um <i>browser</i> ao <i>webserver</i> criado pela câmara <i>Gopro</i> que permite o acesso aos ficheiros por ela criados. . . . .	29
3.4	Esquema geral do funcionamento do Sistema de Visão Computacional. . . . .	30
3.5	Esquema explicativo da estrutura de um pixel (a) e da representação de uma escala de cinzentos (b). . . . .	32
3.6	Exemplo de utilização da binarização numa imagem. . . . .	33
3.7	Histograma de contagem de ocorrência de pixels pretos por coluna da imagem. . . . .	34
3.8	Fluxograma representativo da utilização do algoritmo de <i>pinpoint</i> . . . . .	35
3.9	Exemplo de aplicação de <i>laser pinpoint</i> . . . . .	36
3.10	Paços utilizados para a reconstrução do Laser. . . . .	36
3.11	Comparação entre imagem de acordo com o padrão de laser esperado (a) e com possível existência de ruído (b). . . . .	37
3.12	Fluxograma representativo das regras de determinação de intervalos de ruído. . . . .	38

3.13	Exemplo de aplicação do processo de preenchimento de espaços vazios. . . . .	39
4.1	Sensor usado para a captação dos cenários e seus elementos constituintes. . . . .	42
4.2	Esquema representativo da refração da luz laser através da utilização de um prisma cilíndrico. . . . .	43
4.3	Esquema representativo da relação de distâncias existente entre os elementos do sensor. . . . .	43
4.4	Esquema de cálculo da distância de referência. . . . .	44
4.5	Esquema de cálculo para distâncias inferiores à distância de referência. . . . .	45
4.6	Esquema de cálculo para distâncias superiores à distância de referência. . . . .	46
4.7	Maquete utilizada para o estudo da evolução angular da lente da câmara. . . . .	47
4.8	Superfície que aproxima a evolução angular da lente da câmara em relação às coordenadas dos pixels da imagem. . . . .	48
4.9	Imagem sem tratamento obtida no teste de percepção de profundidade de uma linha laser. . . . .	51
4.10	Resultado do processamento de selecção da linha laser. . . . .	52
4.11	Representação tridimensional de uma linha laser. . . . .	52
4.12	<i>Point-cloud</i> resultante da aquisição de um espaço por rotação do sensor representadas em Perspectiva (a) e Planta (b). . . . .	53

# Lista de Tabelas

3.1	Comparação entre tecnologias de aquisição . . . . .	24
3.2	Especificações da câmera GoPro em modo de Fotografia. . . . .	27
3.3	Alguns comandos usados para controlar a câmara <i>GoPro</i> via <i>WiFi</i> . . . . .	29
4.1	Resultados do Teste de Calibração . . . . .	50



# Lista de Símbolos e Acrónimos

**AD** *Analog-to-Digital*

**CCD** *Charge-Coupled Device*

**CDC** Controlo Directo por Computador

**CMM** *Coordinate-Measuring-Machine*

**IP** *Internet Protocol*

**PS** *Phase Shifting*(Comparação de Fase)

**RGB** *Red Green and Blue* (vermelho, verde e azul)

**ToF** *Time-of-Flight*(Tempo-de-Voo)

**URL** *Uniform Resource Locator*



# CAPÍTULO 1

## Introdução

---

1.1	Enquadramento . . . . .	1
1.2	Motivação . . . . .	2
1.3	Objectivos . . . . .	3
1.4	Sinopse . . . . .	3

---

### 1.1 Enquadramento

Vivemos neste momento na era da informação [2], também conhecida como a era digital. Enquanto por um lado podemos verificar a crescente facilidade de acesso à informação também se tem verificado cada vez mais a digitalização do nosso mundo.

Vê-se, nos dias de hoje, a crescente necessidade do ser humano de trazer o mundo real para o virtual. Uma prova desta necessidade é o projecto de dimensão mundial lançado pela *Google*, a 25 de agosto de 2007 [9], *Google Street View*, que se propunha fornecer aos utilizadores uma visão panorâmica de várias posições ao longo de várias cidades dos Estados Unidos criando assim uma representação digital do que o utilizador poderia observar ao longo dessas ruas. Na figura 1.1 está representada a actual cobertura do *Street View*.

## 1. INTRODUÇÃO

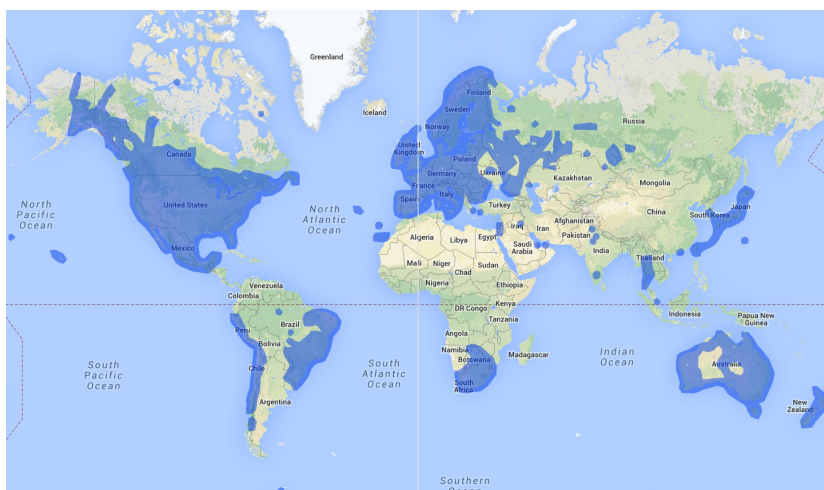


Figure 1.1: Actual cobertura da aplicação "Google Street View" no mundo.

Por existir um crescente interesse em trazer o mundo real para o mundo digital, a modelação tridimensional tem vindo a ser um método cada vez mais requisitado, de forma a criar representações da realidade do nosso mundo, seja ela a modelação tridimensional de partículas microscópicas, pequenos objectos, edifícios ou até mesmo de cidades inteiras.

Embora a modelação tridimensional seja um método bastante requisitado, a sua utilização implica habilidade, conhecimento e necessita tempo proporcional ao detalhe do que se quer modelar, sendo, na maioria dos casos, um processo moroso e exigente. O desenvolvimento de *scanners* 3D tem vindo a crescer como forma de facilitar o processo de aquisição dos dados espaciais necessários para a construção dos modelos tridimensionais.

O *scan* tridimensional, nomeadamente o *scan* laser, tem vindo a ser alvo de pesquisa nos últimos anos de forma a encontrarem-se métodos que auxiliem a criação de modelos 3D. Um exemplo actual da utilização deste tipo de tecnologia é o projecto *Heritage3D* [1](fundado em 2009), projecto patrocinado pelo programa *English Heritage's National Heritage Protection Commissions*, tendo sido cedido o seu desenvolvimento à *School of Civil Engineering and Geosciences, Newcastle University*.

O *scan* 3D é utilizado por muitas áreas uma vez que a facilidade de aquisição de modelos digitais criados pelo *scanner* simplificam e agilizam tarefas específicas. Áreas como Engenharia Civil, Indústria de Construção, Design, Entretenimento, Medicina e Arqueologia são apenas alguns exemplos de áreas que já adoptaram a utilização de *scanners* tridimensionais de forma a simplificarem determinadas tarefas.

## 1.2 Motivação

No âmbito desta área curricular de dissertação de mestrado tendo como motivação a crescente necessidade de utilização deste tipo de tecnologia e visto que esta está em constante evolução, foi



considerado interessante avaliar a possibilidade de criação de um sistema de aquisição perceptual, com a capacidade de reproduzir o modelo tridimensional de zonas interiores. Nestes cenários fechados poderemos considerar as divisões de um edifício, os seus espaços interiores e os objetos neles existentes. Esta foi a motivação para o desenvolvimento deste trabalho na busca de soluções integradas de hardware e software que permitissem a aquisição e tratamento de dados para a percepção deste tipo de modelos tridimensionais.

## 1.3 Objectivos

- Criar um sistema de percepção tridimensional portátil e acessível na sua aplicabilidade;
- Criar uma base de aquisição de dados de percepção tridimensional com a possibilidade de desenvolvimento e potenciação em diferentes áreas de utilização;
- Projectar um sistema com uma relação custo/precisão favorável.

## 1.4 Sinopse

Neste capítulo foi enquadrado o problema existente pela crescente vontade do ser humano trazer o mundo real para o mundo digital, vontade essa demonstrada pelo crescente requisito da utilização de métodos de modelação tridimensionais. Foi também referido quais os motivos e objectivos que levaram à pesquisa e desenvolvimento desta dissertação de mestrado e do sistema de percepção tridimensional proposto.

No capítulo seguinte falar-se-á de algumas das tecnologias e técnicas utilizadas em sistemas já existentes de percepção tridimensional.

## *1. INTRODUÇÃO*

# CAPÍTULO 2

## Estado da Arte

---

<b>2.1</b>	<b>Métodos de Aquisição . . . . .</b>	<b>5</b>
2.1.1	Contacto . . . . .	5
2.1.2	Não-Contacto . . . . .	7
<b>2.2</b>	<b>Processamento de Imagem . . . . .</b>	<b>13</b>
2.2.1	Ruído em Imagens Digitais . . . . .	14
2.2.2	Filtros de Remoção de Ruído . . . . .	15
<b>2.3</b>	<b>Segmentação . . . . .</b>	<b>19</b>
<b>2.4</b>	<b>Sinopse . . . . .</b>	<b>21</b>

---

## 2.1 Métodos de Aquisição

O mundo da tecnologia de percepção tridimensional, com o objectivo de adquirir digitalmente o modelo 3D de um objecto, está dividida em duas categorias bem definidas: as tecnologias de contacto e as de não contacto. Por sua vez, a categoria das soluções das tecnologias de não contacto podem ser definidas como activas ou passivas [5]. Esta categorização permite-nos facilmente determinar a base de funcionamento de cada tipo de tecnologia dependente da categoria à qual a tecnologia pertence.

### 2.1.1 Contacto

Os *scanners* 3D de contacto examinam o objecto através do toque físico enquanto o mesmo está em contacto/repouso sobre uma superfície plana.

O mecanismo de scan de contacto normalmente consiste numa de três formas [4]:

- Um sistema com uma plataforma móvel e braço estático posicionado perpendicularmente em

## 2. ESTADO DA ARTE

relação à plataforma que por sua vez se movimenta de forma a criar contacto entre o objecto e o braço. Este tipo de sistema é ideal para a análise de superfícies de perfil liso, ou superfícies curvas convexas simples.

- Um braço articulado com sensores angulares de alta precisão sendo neste caso o próprio braço robótico que procura o contacto com o objecto em estudo. Este tipo de solução é ideal para avaliar fendas ou espaços interiores com pequenas aberturas, como por exemplo, o interior de uma porca.
- Também poderá ser utilizada uma combinação dos dois métodos referidos anteriormente, de forma a mapear objectos grandes, com cavidades interiores ou superfícies sobrepostas.

Um exemplo deste tipo de tecnologia será o típico 3 “bridge” *Coordinate-Measuring-Machine* (CMM) composto por três eixos, X,Y e Z. Estes eixos são ortogonais entre si de acordo com o típico sistema de coordenadas tridimensionais. Cada eixo tem um sistema de escalamento que indica a localização do mesmo. A máquina, por sua vez, irá ler o input da sonda de contacto através do controlo executado por um operador ou programador e, utilizando as 3 coordenadas (X,Y,Z) de cada um desses pontos, determinar o tamanho e posição com uma precisão tipicamente milimétrica. Os CMM são dispositivos normalmente utilizados em processos de manufatura e montagem, com o objectivo de detectar montagens que não cumpriram o objectivo pretendido *a priori*. Gravando de forma precisa as coordenadas de um determinado objecto este dispositivo gera pontos que podem ser analisados utilizando algoritmos de regressão. Estes pontos são obtidos usando uma sonda que pode tanto ser posicionada manualmente por um operador ou automaticamente via Controlo Directo por Computador (CDC). As CMMs que usam CDC têm a capacidade de ser programados para medirem objectos idênticos repetidamente, sendo por isso as CMM uma forma especializada de robótica industrial [4].

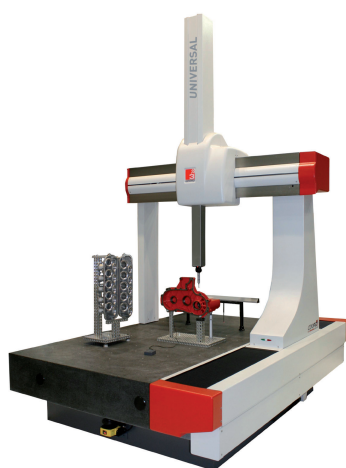


Figura 2.1: Exemplo de um sistema CMM.

### 2.1.2 Não-Contacto

A este grupo pertencem todos os dispositivos de percepção espacial que, tendo como objectivo obter o modelo digital tridimensional do sistema, não recorrem ao contacto físico com os objectos ou espaço em estudo para proceder à aquisição de dados, podendo portanto executar essa aquisição à distância. Estes dispositivos podem ser separados em duas principais categorias: passivos ou activos. A principal forma de distinção entre estas categorias é a forma usada para perceber o sistema. Enquanto que os sensores passivos captam imagens do sistema utilizando a luz já nele existente, os sensores activos distinguem-se pela utilização de projecções especiais de padrões luminosos no sistema, de forma a perceberem o sistema em estudo [5].

#### 2.1.2.1 Passivo

Tal como foi referido anteriormente, este tipo de solução não recorre a qualquer tipo de emissão de radiação, dependendo por isso da reflexão da radiação do ambiente na sua aquisição. Embora na maioria das soluções os dispositivos recorram a detecção da luz visível, existem algumas soluções que utilizam outros tipos de radiação, tal como radiação infra-vermelha. Estes métodos têm a grande vantagem de poderem ser baratos pois não recorrem a hardware especializado, necessitando apenas da utilização de câmaras digitais.

Alguns exemplos de métodos utilizados neste grupo são [5]:

**Estereoscopia** - Este tipo de sistemas usa normalmente um conjunto de duas câmaras desfasadas e apontadas para o mesmo cenário. Através da comparação entre as imagens captadas por cada uma das câmaras, é possível determinar a distância de cada ponto na imagem. Este método é baseado no mesmo princípio de visão estereoscópica que o cérebro humano utiliza para perceber a profundidade e as distâncias.

**Fotometria** - Estes sistemas utilizam apenas uma câmara e retiram informação das distâncias através de múltiplas imagens tiradas de um cenário, através da variação das condições de iluminação do mesmo.

**Técnicas de Silhueta** - Esta técnica utiliza os contornos criados por uma sequência de fotografias em torno de um objecto tridimensional, através de um fundo com grande contraste. Estas silhuetas são estruturadas de forma a formarem uma representação geométrica aproximada do objecto. Uma vez que esta técnica utiliza as extremidades do objecto para criar a sua representação digital, concavidades e outro tipo de detalhes de relevo presentes no objecto tornam-se indetetáveis.

### 2.1.2.2 Activo

Neste tipo de detecção, geralmente, quanto maior o objecto, menor será a precisão e a resolução que podem ser realisticamente alcançadas. Tal como referido anteriormente, este tipo de *scanners* utiliza algum tipo de projecção de radiação de forma a executar a percepção do cenário, sendo maioritariamente utilizada a projecção de lasers para proceder a essa percepção. Os *scanners* laser podem operar com base em um de três princípios: triangulação, *Time-of-Flight* (Tempo-de-Voo) (ToF) e *Phase Shifting* (Comparação de Fase) (PS) [1].

#### Triangulação

Os *scanners* 3D que utilizam este princípio calculam as suas coordenadas triangulando a posição de um ponto, ou mesmo a posição de uma linha laser. Alguns sistemas de triangulação requerem que o objecto seja colocado numa mesa giratória de forma a rodar o objecto em estudo em frente a um *scanner* estático. De forma alternativa, alguns sistemas utilizam o *scanner* montado num braço mecânico. Tipicamente, neste tipo de sistemas, a distância entre o objecto e o *scanner* é inferior a 1m e comumente obtêm precisão de medição da ordem dos 0.1mm [1].

Embora não proporcionem o alto nível de precisão associada aos laser que utilizam montagens com braços mecânicos, existem também sistemas de triangulação que permitem, através de sistemas de prismas ou espelhos, realizar a percepção automaticamente. Estes sistemas podem ser montados num típico tripé de fotografia, de forma a obterem imagens tridimensionais do sistema em estudo, a distâncias até 2m. Este tipo de sistema tem tendência a ter um *design* mais portátil sendo ideal na captura de pequenas características arquitetónicas, como por exemplo, esculturas detalhadas. Por fim, alguns sistemas baseados em triangulação permitem medições a distâncias até 25m, porém, a esta distância é esperada a degradação da precisão. Este tipo de solução normalmente apresenta baixa performance quando avaliando cenários que apresentem uma iluminação solar muito forte sendo necessário em algumas situações fornecer sombra na zona que está a ser percepcionada, de forma a obter uma melhor medição [1].

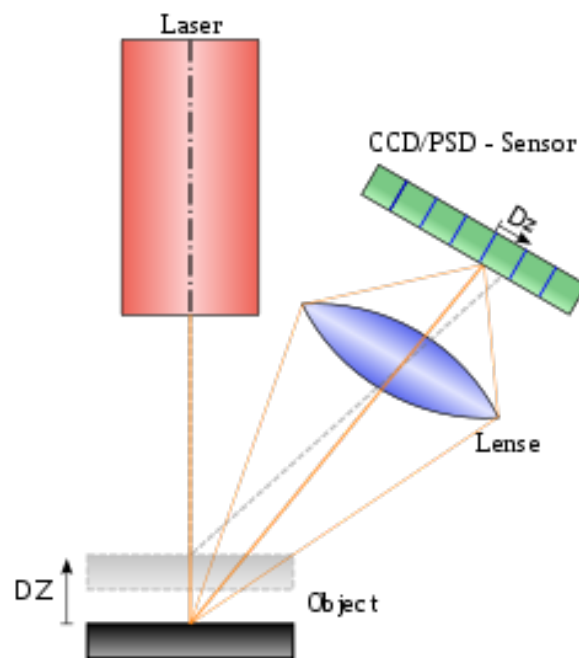


Figura 2.2: Esquema de funcionamento de um *scanner* 3D utilizando triangulação.

De forma a proporcionar uma melhor compreensão do funcionamento da triangulação (figura 2.2) como método de aquisição de dados, mais à frente será exemplificado um modelo matemático no qual este tipo de tecnologia baseia o seu funcionamento, de forma a adquirir os dados de percepção do cenário em estudo. Embora seja na maioria dos casos de utilização deste método projectado um tipo de padrão laser, tal como os apresentados na figura 2.3, para facilitar explicação será exemplificada a determinação da distância para um ponto.

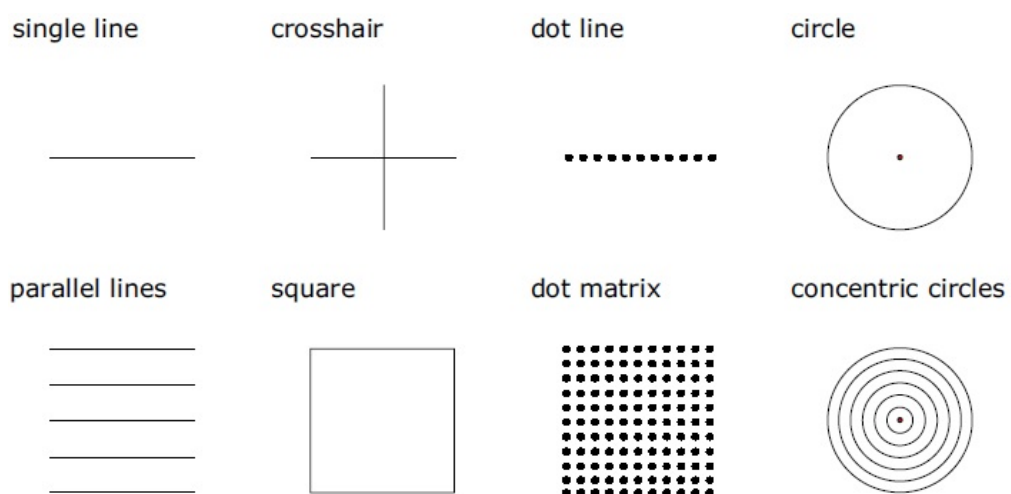


Figura 2.3: Alguns exemplos de padrões luminosos utilizados

## 2. ESTADO DA ARTE

A figura 2.4, demonstra a relação entre a distância existente do sensor  $C$  ao ponto incidência do laser num objecto  $k$ , relativamente a um plano de referência e a sua disparidade  $d$ . De forma a expressar as coordenadas 3D do objecto em pontos, é necessário considerar o sistema de coordenadas com origem no centro de perspectiva do sensor (ponto  $C$ ). O eixo  $Z$  é ortogonal relativamente ao plano da imagem orientada para o objecto, o Eixo  $X$  é perpendicular ao eixo  $Z$  na direcção da base entre o sensor e o emissor laser, por fim o eixo  $Y$  é ortogonal a  $X$  e  $Z$ , criando um sistema de coordenadas de acordo com a regra da mão direita [14].

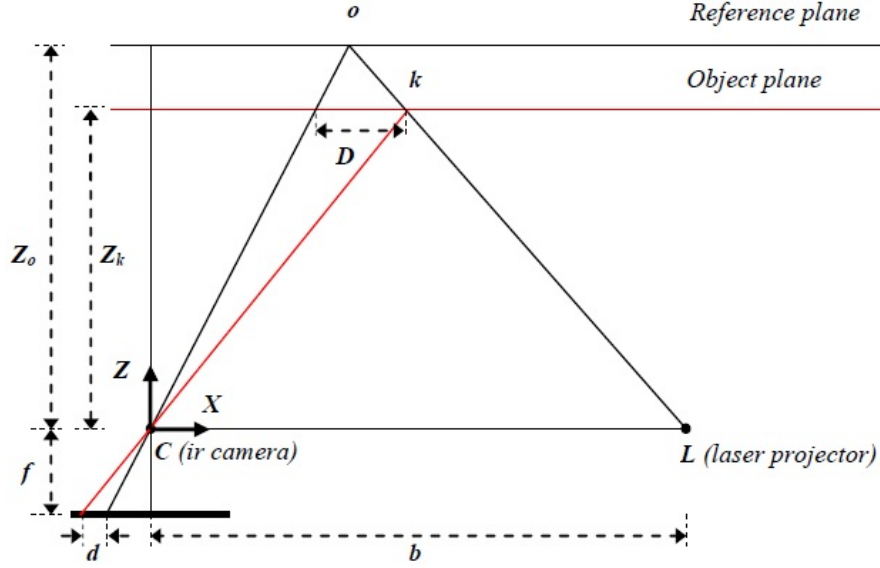


Figura 2.4: Esquema de relação entre a profundidade relativa e a disparidade medida.

Assumindo que um objecto está situado no plano de referência a uma distância  $Z_0$  do sensor, e o ponto laser incidente no objecto  $o$  é capturado na imagem pelo sensor, se o objecto for aproximado ou afastado do sensor, a localização do ponto laser na imagem será desfasada em  $X$ . Este desfasamento é medido no espaço da imagem como a disparidade  $d$ , correspondente a um ponto  $k$  no espaço do objecto.

A partir da semelhança de triângulos obtemos [14, 3]:

$$\frac{D}{b} = \frac{Z_0 - Z_k}{Z_0} \quad (2.1)$$

e:

$$\frac{d}{f} = \frac{D}{Z_k} \quad (2.2)$$

onde  $Z_k$  representa a distância (profundidade) do ponto  $k$  no espaço do objecto,  $b$  é o tamanho da base de distância entre a câmara e o emissor de laser,  $f$  é a distância focal do sensor,  $D$  é o desfasamento



do ponto  $k$  no espaço do objecto e  $d$  é o desvio observado no espaço da imagem. Substituindo  $D$  na equação 2.2 e resolvendo a mesma em ordem a  $Z_k$  obtemos:

$$Z_k = \frac{Z_0}{1 + \left(\frac{Z_0}{fb}\right)^d} \quad (2.3)$$

A equação 2.3 representa o modelo matemático básico do cálculo da profundidade a partir da disparidade observada na imagem e das constantes  $Z_0$ ,  $f$ ,  $b$  que podem ser determinadas a partir da calibração do sensor [14]. As coordenadas de cada ponto do objecto podem então ser calculadas a partir das coordenadas da imagem e a sua escala:

$$X_k = -\frac{Z_k}{f}(x_k - x_o + \delta_x) \quad Y_k = -\frac{Z_k}{f}(y_k - y_o + \delta_y) \quad (2.4)$$

onde  $x_k$  e  $y_k$  são as coordenadas do ponto na imagem,  $x_o$  e  $y_o$  as coordenadas do ponto de referência e  $\delta_x$  e  $\delta_y$  os coeficientes de correcção da distorção da lente do sensor [14].

Um possível exemplo de um sensor que utiliza este tipo de metodologia para a realização da aquisição tridimensional é o sensor *Microsoft Kinect* [13].

### Time-of-Flight (tempo de voo) - ToF

Este tipo de sistema é baseado no conceito de avaliar as distâncias através do tempo de voo de um pulso de laser, sendo a sua utilização apropriada em actividades como preservação arquitetónica ou arqueológica, apresentando normalmente uma precisão entre 3-6mm. Através da medição do tempo de viagem de ida-e-volta de um pulso de energia laser, estes sistemas conseguem avaliar a distância entre o *scanner* e os objectos do sistema em estudo. Tipicamente este tipo de sistemas tem um intervalo de funcionamento de 2-300m sendo esperado que tenham a capacidade de adquirir a cada minuto dezenas de milhares de pontos através da deflecção do pulso de laser na superfície dos objectos utilizando um espelho rotativo ou prisma.

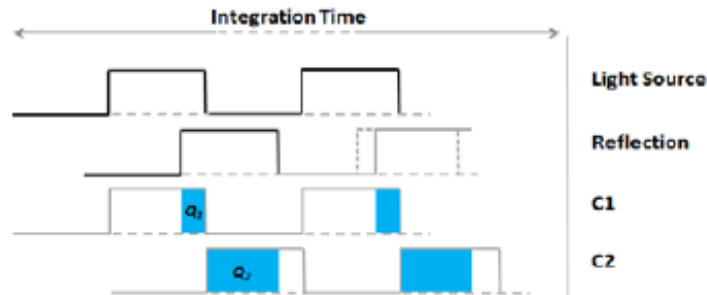


Figura 2.5: Exemplo de dados utilizados para determinação de distâncias usando ToF.

Neste método, a forma de perceber as distâncias (profundidade) do sistema é realizada a partir da avaliação da energia luminosa que, sendo primeiro emitida por uma fonte luminosa durante um

breve período  $\Delta t$ , é depois reflectida pelas superfícies do cenário, sendo mais tarde captada a sua reflexão em cada pixel do sensor. Em paralelo, são utilizadas duas janelas fora de fase,  $C_1$  e  $C_2$ , com o mesmo  $\Delta t$ . Nessas janelas são acumuladas cargas eléctricas durante as amostragens,  $Q_1$  e  $Q_2$ , sendo depois medidas e usadas de forma a determinar as distâncias entre os objectos e o sensor a utilizando a fórmula 2.5 [15]:

$$d = \frac{1}{2}c\Delta t \left( \frac{Q_2}{Q_1 + Q_2} \right) \quad (2.5)$$

Onde  $c$  é a constante que representa a velocidade da luz ( $3.0 \times 10^8 m/s$ ).

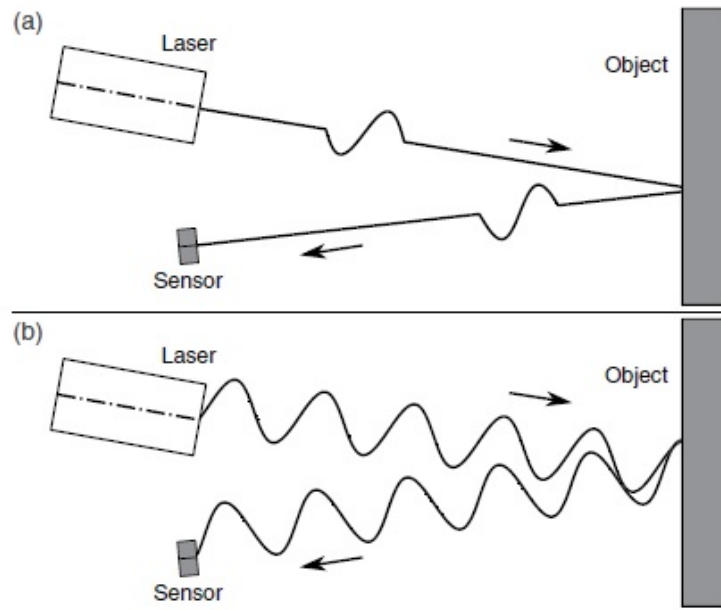


Figura 2.6: Representação dos métodos de aquisição de distâncias ToF(a) e PS(b).

### Phase-Shifting (comparação de fase) - PS

Apesar deste tipo de sistemas apresentarem um método de funcionamento e uma precisão semelhantes aos oferecidos pelos sistemas que utilizam ToF, os sistemas PS funcionam de uma forma ligeiramente diferente (figura 2.6). Este tipo de *scanner* baseia a medição das distâncias do cenário usando as diferenças apresentadas entre o sinal emitido e o sinal que é recebido por reflexão. Estes sistemas apresentam uma maior quantidade de dados de captura (cerca de milhões de pontos por minuto), resultando numa *point-cloud* com densidade bastante superior. Uma elevada densidade de pontos, normalmente exige elevados requisitos de hardware, de forma a ser possível o processamento dessa grande quantidade de informação.

Em contraste com o método utilizado em ToF que utiliza pulsos, o PS usa uma onda modulada luminosa contínua, tipicamente uma onda quadrada ou sinusoidal e retira múltiplas amostras por medição sendo cada uma dessas amostras desfasada em  $90^\circ$ , perfazendo um total de quatro amostras. Usando esta

técnica, a fase do ângulo existente entre a radiação luminosa emitida e a sua reflexão,  $\varphi$ , e a distância,  $d$ , podem ser calculadas através das fórmulas 2.6 e 2.7 [15]:

$$\varphi = \arctan\left(\frac{Q_3 - Q_4}{Q_2 - Q_1}\right) \quad (2.6)$$

$$d = \frac{c}{4\pi f} \varphi \quad (2.7)$$

Segue-se que a intensidade do pixel ( $A$ ) e o *offset* ( $B$ ) podem ser calculados por:

$$A = \frac{\sqrt{(Q_1 - Q_2)^2 + (Q_3 - Q_4)^2}}{2} \quad (2.8)$$

$$B = \frac{Q_1 + Q_2 + Q_3 + Q_4}{4} \quad (2.9)$$

Na figura 2.7 está representado um exemplo dos dados normalmente obtidos para a realização dos cálculos acima apresentados, nomeadamente as 4 janelas,  $C_n$  e as suas respectivas amostragens,  $Q_n$ .

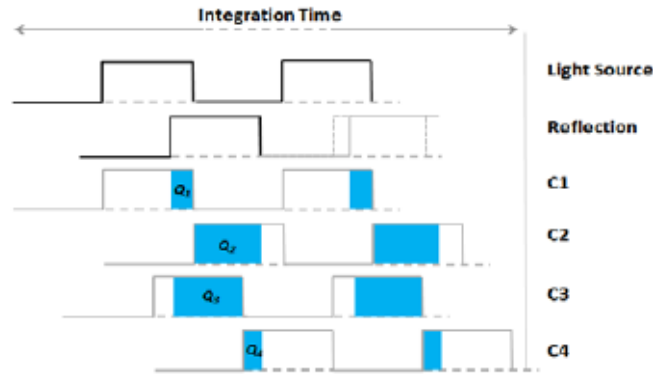


Figura 2.7: Exemplo de dados utilizados para determinação de distâncias usando PS.

## 2.2 Processamento de Imagem

Uma grande quantidade dos métodos de aquisição de não-contacto referidos anteriormente, como por exemplo o método de triangulação, são métodos que utilizam como base sensorial imagens digitais captadas a partir de câmaras digitais. A imagem em si, nestes casos, não fornece directamente os dados da profundidade do sistema. É, por esta razão, necessário recorrer ao processamento dessas imagens de forma a tornar possível a extracção da informação perceptual nelas existente. Embora

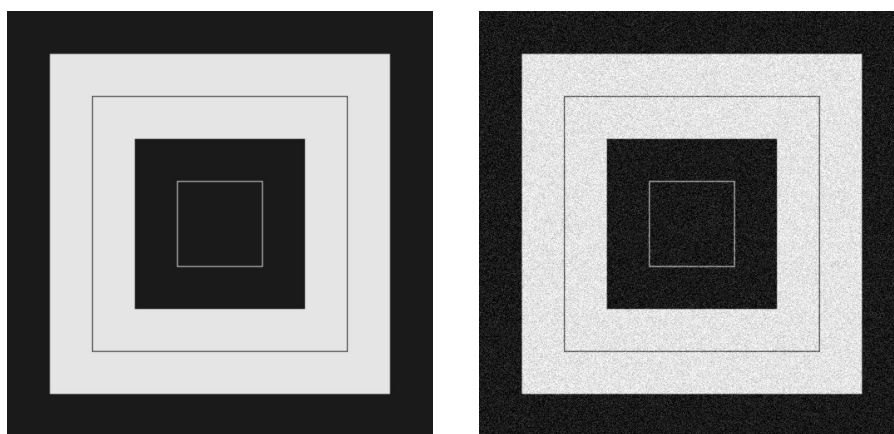
cada método implique a utilização de diferentes tipos de processamento de imagem há dois elementos aos quais a maioria terá de ter em consideração: o ruído existente na imagem e a capacidade de segmentação da imagem.

### 2.2.1 Ruído em Imagens Digitais

É dado o nome de ruído a qualquer degradação patente na imagem podendo esta ser produzida pelo sensor, meio de transmissão dos dados da imagem ou qualquer outro factor que introduza erros na imagem. Por outras palavras, é possível afirmar que o ruído é o resultado dos erros causados pelo processo de aquisição de imagem que resulta na criação de valores de pixéis que não reflectem a verdadeira intensidade da cena real [16]. O ruído presente nas imagens pode ser classificado como:

#### Ruído de Amplificação (Gaussiano)

Como nas câmaras a cores, normalmente o canal de cor azul utiliza maior amplificação que nos restantes canais (verde e vermelho), é possível verificar uma quantidade de ruído superior no referido canal. O ruído de amplificação é maioritariamente constituído por ruído proveniente do sensor, apresentando-se como um nível constante de ruído nas áreas escuras da imagem. No ruído Gaussiano, cada pixel da imagem sofrerá normalmente uma pequena alteração em relação ao seu valor original. Um histograma, gráfico que representa a quantidade de distorção de um pixel em relação à frequência com que o mesmo ocorre, demonstra a distribuição normal do ruído. Enquanto é possível a existência de outras distribuições, a distribuição Gaussiana normal é geralmente um bom modelo de aproximação quando tendo em consideração o teorema de limite central que diz que a soma de diferentes ruídos tende a aproximar-se de uma distribuição Gaussiana [16].



(a) Imagem sem ruído Gaussiano.

(b) Imagem com ruído Gaussiano.

Figura 2.8: Comparação entre duas imagens com e sem ruído Gaussiano.

### Ruído *Salt-n-Pepper* (Impulsivo)

O ruído *salt-n-pepper* (sal e pimenta) também conhecido como ruído impulsivo, ruído de pico, ruído aleatório ou ruído independente, é definido pela existência de pixéis na imagem que diferem de forma muito significativa dos pixéis vizinhos, quer seja em cor ou na sua intensidade. Este tipo de degradação pode ser causado por funções de *sharpening* (funções que alteram a nitidez da fotografia) ou perturbações súbitas no sinal da imagem. Geralmente este tipo de ruído apenas afecta uma pequena quantidade de pixéis na imagem. Quando observada, a imagem contém pontos pretos e brancos, daí a designação de ruído *salt-n-pepper* (sal e pimenta). Outras fontes deste tipo de ruído podem ser criadas pela existência de partículas de pó dentro da câmara, sobreaquecimento ou defeitos nos elementos constituintes do *Charge-Coupled Device* (CCD). Imagens que apresentem este tipo de ruído terão pixéis escuros em zonas brilhantes e pixéis brancos em zonas escuras. Este tipo de ruído pode ainda ser causado pela existência de pixéis mortos no sensor, erros no conversor *Analog-to-Digital* (AD) e erros na transmissão dos bits.

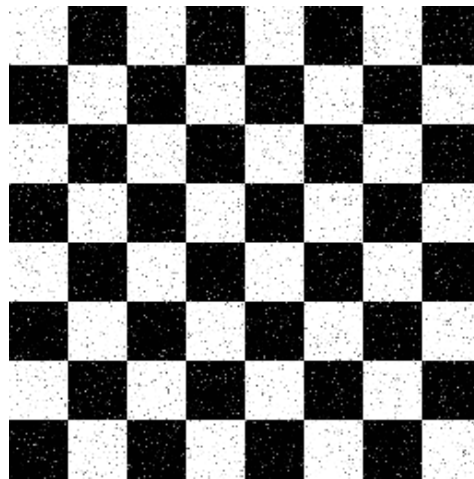


Figura 2.9: Imagem afectada pelo ruído *Salt-n-pepper*.

#### 2.2.2 Filtros de Remoção de Ruído

Quando uma imagem é capturada por uma câmara ou por outro sistema de aquisição de imagem é habitual os sistemas de visão computacionais não terem a capacidade de usar directamente a imagem capturada para executarem as funções para os quais foram projectados. A imagem capturada pode estar corrompida por variações aleatórias de intensidade, iluminação ou terem pouco contraste que deve ser tratado nas fases iniciais do processamento de visão computacional. Tal como foi abordado na secção anterior, essas perturbações são consideradas como ruído presente nas imagens. Os filtros são métodos de processamento de imagem com o objectivo de eliminar ou atenuar estas características

## 2. ESTADO DA ARTE

indesejadas nas imagens capturadas.

Em processamento de imagem, tal como referido por R.Gonzalez [8], são executadas operações de vizinhança que trabalham com o valor do nível de cinzento dos pixéis da vizinhança e os valores correspondentes de uma sub-imagem que tem dimensão igual à dimensão da dita vizinhança. A terminologia normalmente utilizada para referenciar essa sub-imagem varia comumente entre filtro, máscara, *kernel*. Os valores na sub-imagem do filtro são chamados de coeficientes e não pixéis.

Apesar de a filtragem ter as suas raízes na utilização da transformada de *Fourier* para processamento de sinal no chamado domínio da frequência, o tipo de operações de filtragem que serão referidas de seguida são operações efectuadas directamente nos pixéis da imagem, dessa forma, para haver uma diferenciação entre a filtragem executada no domínio da frequência e a filtragem que se irá tratar, utilizar-se-á o termo filtragem espacial para nos referirmos a este tipo de processo.

A mecânica de utilização de filtragem espacial está ilustrada na figura 2.10. O processo consiste simplesmente na movimentação da máscara do filtro de convolução ponto a ponto na imagem. Em cada ponto (x,y), a resposta do filtro nesse ponto é calculada usando a relação predefinida pelo mesmo.

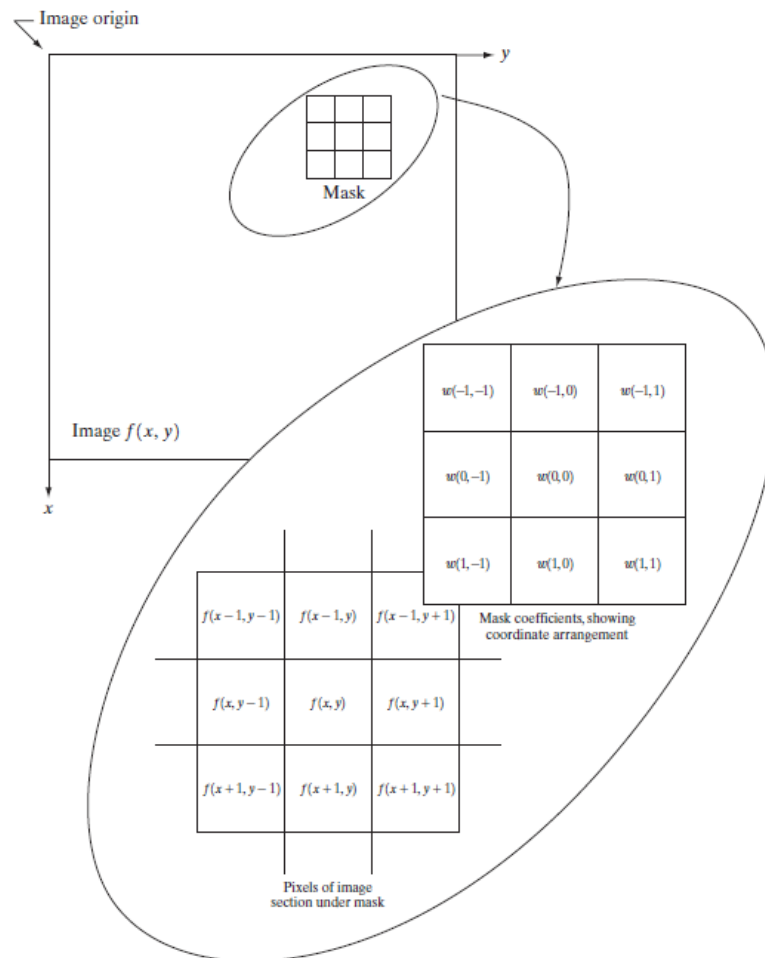


Figura 2.10: Mecânica de funcionamento de filtragem espacial usando uma máscara de tamanho 3x3.

### 2.2.2.1 Filtros Lineares

Para filtragem espacial Linear, a resposta é dada pela soma do produto dos coeficientes do filtro com valores dos pixels da imagem presentes na área definida pela máscara do filtro. Para um filtro 3x3 tal como o representado na figura 2.10, o novo valor que o pixel central  $f(x,y)$  tomará será definido pela seguinte equação [8]:

$$R = w(-1, -1)f(x - 1, y - 1) + w(-1, 0)f(x - 1, y) + \dots + w(0, 0)f(x, y) + \dots + w(1, 0)f(x + 1, y) + w(1, 1)f(x + 1, y + 1) \quad (2.10)$$

Como é possível verificar na fórmula 2.10, o valor que o pixel central tomará,  $R$ , é o resultado da soma dos produtos dos coeficientes da máscara pelos valores dos pixels aos quais a máscara está sobreposta. Tendo em consideração a fórmula e a aplicação de um filtro linear numa imagem  $f$  de tamanho  $M \times N$  e aplicando uma máscara de tamanho  $m \times n$ , o valor de cada pixel da imagem poderá ser calculado pela expressão [8]:

$$g(x, y) = \frac{1}{\sum w(s, t)} \sum_{s=-a}^a \sum_{t=-b}^b w(s, t)f(x + s, y + t) \quad (2.11)$$

De forma a se filtrar a imagem na sua totalidade esta equação necessita ser aplicada a todos os pontos da imagem pertencentes aos intervalos: em  $x \in [0, M-1]$  e  $y \in [0, N-1]$ . Sendo desta forma assegurado que todos os pixels da imagem são processados. Na figura 2.11 está representado um exemplo de aplicação de um filtro de média encontrando-se à esquerda a imagem original e à direita a imagem resultante da aplicação do filtro.

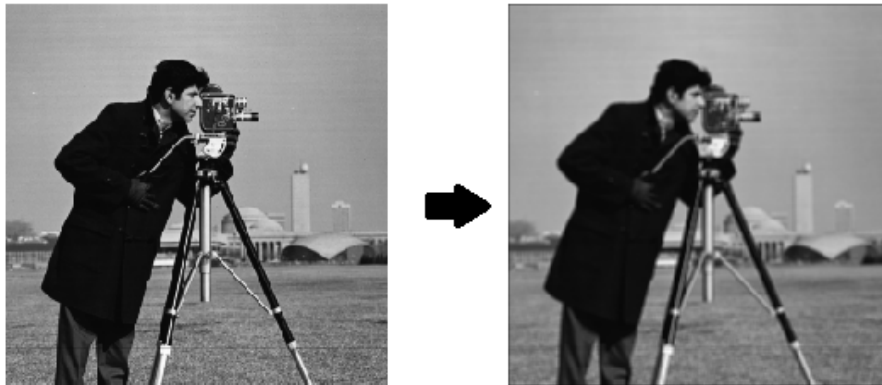


Figura 2.11: Exemplo de aplicação de um filtro de média linear.

### 2.2.2.2 Filtros não-Lineares

Os filtros espaciais não lineares, à semelhança dos filtros referidos anteriormente neste capítulo, também operam com base em vizinhanças e as mecânicas de movimento da máscara do filtro ao longo da imagem. No entanto, este tipo de filtro baseia o seu funcionamento em operações de condições sobre o valor dos pixels da vizinhança em estudo, ordenando-os e finalmente substituindo o pixel central da máscara pelo valor do pixel determinado pela função lógica de avaliação. O exemplo mais conhecido desta categoria de filtros é o filtro de mediana, que, tal como o nome implica, substitui o valor do pixel pela mediana dos níveis de cinzento pertencentes à vizinhança desse pixel. Os filtros de mediana são tão populares por fornecerem, para certos tipos de ruído aleatório, uma excelente capacidade de redução de ruído sem desfocarem tanto a imagem quando comparado com a maioria dos filtros lineares de tamanho semelhante. Estes filtros são particularmente efectivos em imagens que apresentem ruído impulsivo.

A mediana,  $\tilde{x}$ , de um determinado conjunto de valores é determinada de maneira a que metade dos valores seja menor ou igual a  $\tilde{x}$  e a outra metade maior ou igual que  $\tilde{x}$ . Para se realizar uma operação de filtragem usando um filtro de mediana num ponto de uma imagem, primeiro os valores do pixel em questão e dos pixels da sua vizinhança são ordenados de acordo com a sua intensidade, depois, dentro desse conjunto é determinada a mediana e atribuído esse valor ao pixel. Numa vizinhança de 3 x 3, a mediana será o quinto maior valor do intervalo, numa vizinhança de 5 x 5 o décimo terceiro, etc. Na figura 2.12 está apresentado um exemplo de aplicação de um filtro não linear de mediana com uma máscara de tamanho 3x3.

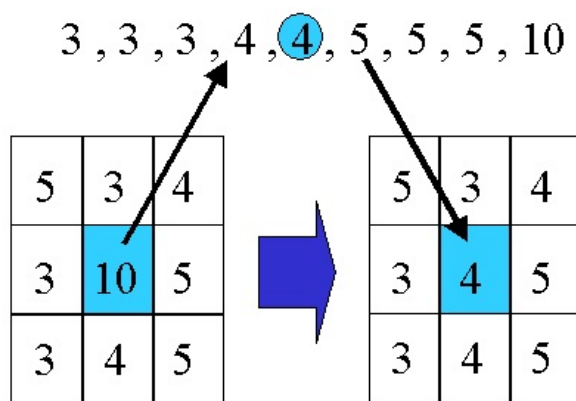


Figura 2.12: Exemplo de aplicação de um filtro de mediana com uma máscara 3x3.

Assim sendo, a principal função dos filtros de mediana é forçar pontos com valores de cinzento distintos tornarem-se mais parecidos com a sua vizinhança. De facto, conjuntos isolados de pixels escuros ou claros em relação ao resto da sua vizinhança, com área menor a  $\frac{n^2}{2}$  (metade da área filtrada), são eliminados por um filtro de mediana de tamanho  $n \times n$ . Quando se utiliza o termo "eliminar", quer-se pois referir ao processo de alteração do valor de intensidade do pixel para a sua



mediana. Essa eliminação faz com que os pixels isolados anteriormente referidos deixem de se destacar.

## 2.3 Segmentação

Sob o ponto de vista computacional, é chamada à segmentação o processo de partição de uma imagem digital e múltiplos segmentos. O objectivo da segmentação é o de simplificar a imagem alterando a sua representação de forma a facilitar a sua análise. Este tipo de ferramenta é tipicamente utilizado para localizar objectos e limites tais como linhas, curvas, etc. De uma forma mais precisa, a segmentação de imagens é o processo de atribuição de uma categoria a cada pixel de uma imagem, de forma a que pixels da mesma categoria partilhem determinadas características. O resultado da segmentação de uma imagem será um conjunto de segmentos que ocupam, no seu conjunto total, toda a imagem ou um conjunto de contornos extraídos da imagem. Cada um dos pixels na região da imagem apresentam semelhanças em respeito a alguma característica ou propriedade computacional, como a cor, intensidade ou textura. Regiões adjacentes são significativamente diferentes quando comparadas essa(s) determinada(s) característica(s).

### *Thresholding*

É o método mais simples de segmentação, baseando-se num nível, isto é um limite, para transformar uma imagem de escala de cinzentos numa imagem binária. O ponto chave deste método é a selecção do valor de *threshold*. Na prática, são realizados séries de testes onde se examinam as imagens segmentadas através deste método enquanto se ajusta o valor de *threshold*, sendo o melhor valor definido com base na observação da imagem binarizada que melhor separa os detalhes pretendidos (podendo estes testes ser evitados quando utilizados métodos para aquisição automática do valor de *threshold* como por exemplo o método de Otso) [6]. Esta é uma técnica muito usada em aplicações industriais, especialmente quando se quer examinar detalhes particulares de um objecto.

As técnicas de determinação do valor de *threshold* frequentemente envolvem a análise de histogramas dos níveis de intensidade das imagens (figura 2.13).

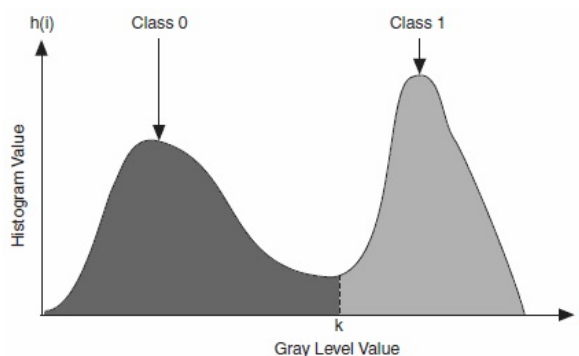


Figura 2.13: Exemplo de análise de histograma para determinação de valor de *threshold*.

Na figura 2.13 está demonstrado um exemplo de análise de um histograma de intensidade de cinzentos de forma a utilizar o método de *Thresholding*, em que  $k$  representa o valor de *threshold* escolhido de forma a melhor separar os objectos escuros, à esquerda de  $k$ , do fundo claro, à direita de  $k$ .

### ***Edge Detection***

A segmentação tem sido abordada como a forma de encontrar regiões de uniformidade nas imagens, como por exemplo a superfície de um objecto. *Edge Detection* é um método alternativo para realizar a segmentação de uma imagem e consiste na detecção de limites nos objectos da imagem, isto é, locais da imagem onde ocorrem diferenças significantes na intensidade dos pixels da imagem. O método de *Edge detection* tem sofrido uma grande evolução ao longo das últimas décadas e é possível distinguir dois tipos distintos de utilização deste método: *Template Matching* (TM) e *Differential Gradient* (DG). Em ambos os casos, o objectivo deste método consiste na detecção dos locais onde a intensidade de gradiente,  $g$ , é suficientemente grande para ser definida como indicador da existência de um limite de um objecto. Com a ajuda de máscaras apropriadas, os operadores DG e TM, estimam o valor da intensidade do gradiente local. No caso do operador do tipo DG, apenas 2 máscaras são necessárias (para a direcção  $x$  e  $y$ ). No caso do TM, é normal usar-se até 12 máscaras capazes de estimar componentes locais do gradiente em diferentes direcções [6].

Algumas das conhecidas máscaras utilizadas em *Edge Detection* são:

#### **Operador de Roberts 2x2:**

$$R_{x'} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad R_{y'} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2.12)$$

#### **Operador de Sobel 3x3:**

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2.13)$$

#### **Operador de Prewitt 3x3:**

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad S_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (2.14)$$

Na figura 2.14 é possível observar o resultado de uma segmentação realizada utilizando *edge detection*.



Figura 2.14: Exemplo de utilização de *edge detection*.

## 2.4 Sinopse

Neste capítulo, foram abordadas tecnologias e métodos utilizados pela maioria dos sistemas de percepção tridimensional de forma a possibilitarem a aquisição dos dados espaciais necessários para a reprodução dos modelos tridimensionais desejados. Estes métodos e tecnologias formam o *Estado da Arte* deste tipo de sistemas que podem ser distinguidos pelo tipo de método de aquisição que utilizam. Sendo a maioria destes métodos de aquisição sistemas de *não-contacto*, foi também importante incluir no estudo do *Estado da Arte* alguns dos principais métodos utilizados em *Visão Computacional* pois estes são muitas vezes necessários como métodos de extracção de dados de informação.

No capítulo seguinte, com base no estudo bibliográfico presente neste capítulo, será proposto um processo que permita o cumprimento dos objectivos traçados no capítulo anterior.

## *2. ESTADO DA ARTE*

# CAPÍTULO 3

## Proposta

---

<b>3.1</b>	<b>Processo . . . . .</b>	<b>23</b>
<b>3.2</b>	<b>Hardware . . . . .</b>	<b>26</b>
<b>3.3</b>	<b>Software . . . . .</b>	<b>28</b>
3.3.1	Sistema de Comunicação . . . . .	28
3.3.2	Sistema de Visão Computacional . . . . .	30
3.3.3	Sistema de Cálculo . . . . .	40
<b>3.4</b>	<b>Sinopse . . . . .</b>	<b>40</b>

---

### 3.1 Processo

Sendo o objectivo principal desta dissertação a criação de uma plataforma de apoio à modelação tridimensional de espaços fechados, é necessário determinar as características desse tipo de espaços e quais os requisitos de modelação pretendidos de forma a se escolher a tecnologia de aquisição a utilizar no processo. As principais características que se esperam observar nesses espaços fechado são: a iluminação do espaço, natural ou artificial, ter a possibilidade de ser controlada e a existência de limites físicos em todas as direcções de forma a que a distância entre o sensor e as superfícies do espaço não ultrapasse os 5 metros. Sendo assim, de forma a se poder escolher um método de aquisição adequado, é requerido que este seja portátil, tenha uma precisão o mais elevada possível e um alcance até 5 metros. Tendo em consideração o estado da arte presente nesta dissertação (capítulo 2) e tendo em conta que um dos requisitos da tecnologia de aquisição para este projecto seja esta ter facilidade na sua portatibilidade as tecnologias de contacto serão excluídas. Como há a possibilidade do espaço fechado em avaliação não estar iluminado, tendo em consideração a pesquisa realizada na secção 2.1.2 do estado da arte é possível concluir que as tecnologias de não contacto passivas, por utilizarem a iluminação já existente no espaço em estudo para realizarem a sua percepção, também não se adequam

### 3. PROPOSTA

como solução para a abordagem de processo pretendida.

Resta-nos por essa razão proceder à selecção do sistema de aquisição, de entre as tecnologias oferecidas pelos *scanners* de não contacto activos, também conhecidos como *scanners* laser. Com base no artigo [1], e sabendo que não existe uma tecnologia ideal, mas sim tecnologias que se adequam melhor que outras a determinadas situações, é facilitada a escolha do tipo de *scanner* com base na resposta às seguintes perguntas:

- Qual o output desejado?
- Qual o tamanho do objecto de estudo?
- Qual o nível de precisão desejado?
- Qual a resolução de medição?
- Qual o orçamento?

Tentando responder a estas perguntas, e de acordo com as características do que foi considerado um espaço fechado, é possível determinar os requisitos desejados para o processo. É portanto desejado que o processo seja capaz de obter uma nuvem de pontos (*point-cloud*) com a informação das coordenadas cartesianas (X,Y,Z) de cada ponto de forma a poder ser utilizado para a reconstrução de um modelo tridimensional do cenário em estudo. Para isso é necessário que o sensor tenha a capacidade de realizar uma avaliação de 360° de todo o espaço.

A determinação do tamanho do objecto de estudo pretendido não é exacta, pois este pode tomar formas variadas, mas assumir-se-á que o espaço fechado estará dentro do intervalo de 0.1m a 5m de raio.

Como se tem o objectivo de realizar uma modelação o mais próxima da realidade possível e há a possibilidade da existência de artefactos ou objectos no interior do espaço fechado que apresentem formas detalhadas, é desejado que a qualidade de precisão dos pontos obtidos seja o mais próximo da realidade possível, exigindo um nível de resolução e precisão elevadas por parte da informação obtida pelo sensor.

Sendo um dos objectivos desta dissertação também a tentativa de encontrar uma forma de simplificar a utilização deste tipo de sistemas, tentar-se-á utilizar uma tecnologia que seja acessível para além dos outros requisitos referidos anteriormente. Tendo agora em consideração os requisitos nomeados é necessário proceder à comparação das tecnologias ToF, Triangulação e PS. A tabela em baixo apresentada mostra uma comparação das características deste tipo de tecnologias [17].

Tabela 3.1: Comparação entre tecnologias de aquisição

	Vel. de Aquisição	Tamanho dos objectos	Melhor detalhe	Preço	Portabilidade
Triângulação	Minutos a Horas	1mm-1m	50 $\mu$ m	125-12500€	Sim-Semi
ToF	Minutos	>10 mm	1mm	12500€	Semi
PS	Segundos-Minutos	>10 mm	1mm	12500€	Semi

Observando a tabela 3.1 podemos verificar algumas das características teóricas de cada um destes tipos de tecnologia. Enquanto o PS se destaca pela sua notável capacidade de velocidade de aquisição, derivada da sua considerável capacidade de produção de pontos da sua *point-cloud* aquando da aquisição, a triangulação destaca-se não só nas suas potencialidades de detalhe superiores às outras duas tecnologias, como também a sua maior acessibilidade em termos de recursos económicos e portabilidade. Com isto em mente, sabendo que se quer fornecer uma plataforma de apoio à aquisição de dados 3D que seja acessível a todos, as características da triangulação são as que se consideraram melhor encaixar nos objectivos desta dissertação sendo porém necessário tentar proceder a uma melhoria da capacidade desta tecnologia de perceber objectos maiores, mesmo tendo de se sacrificar alguma da sua precisão.

Tentar-se-á, na Validação (capítulo 4), apresentar uma proposta baseada em triangulação que cumpra as especificações referidas.

### Explicação Geral do funcionamento do Processo

Na imagem 3.1 está apresentado um esquema que demonstra o funcionamento geral do processo proposto.

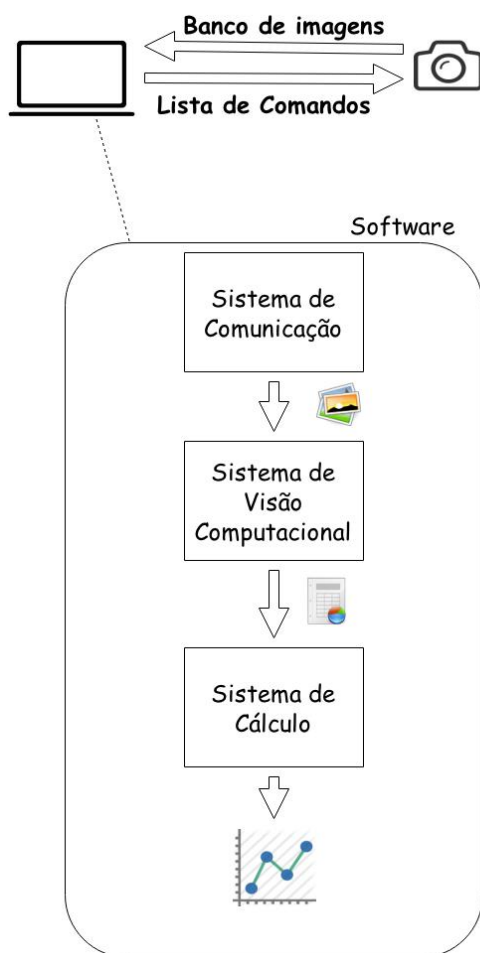


Figura 3.1: Esquema geral do funcionamento do processo proposto.

### 3. PROPOSTA

A partir da base da plataforma no computador, o utilizador envia um comando para o sensor começar a recolha de imagens. Após ter recebido o comando, o sensor começa a captar fotografias do espaço em estudo em intervalos de tempo previamente definidos no comando enviado pelo utilizador. As fotografias vão sendo guardadas numa base de dados à medida que o sensor vai realizando a captura das mesmas. Paralelamente ao processo de aquisição está a correr um processo que vai acedendo ao banco de imagens através do *Sistema de Comunicação* e envia, uma a uma, as imagens ao *Sistema de Visão Computacional* de forma a que este extraia as coordenadas da imagem atravessadas pela linha laser. Durante essa avaliação, o *Sistema de Visão Computacional* inserindo numa tabela a informação extraída da imagem. Após concluída a avaliação de todas as imagens do banco de imagens e ter sido dada como terminada a aquisição do espaço, o *Sistema de Visão Computacional* envia a tabela criada com as coordenadas extraídas da imagem ao *Sistema de Cálculo* que por sua vez, recorrendo a cálculos matemáticos, as interpreta de forma a obter as coordenadas cartesianas (X, Y, Z) do cenário estudado para que seja possível realizar a reprodução de um modelo tridimensional do espaço estudado.

## 3.2 Hardware

A componente hardware da proposta referida na secção anterior desempenha as funções de interacção com o cenário em estudo. Essa interacção baseia-se na projecção de um padrão luminoso de luz laser e a captação de imagens do cenário, de forma a se ir criando uma memorização das zonas pelas quais o laser incidiu sobre as superfícies do cenário, para que mais tarde possam ser interpretadas pela componente de software do processo. A componente de hardware é então constituída pelo sistema físico que compõe o sensor de aquisição espacial. O sensor proposto será portanto constituído por uma câmara digital, um emissor laser e uma base que servirá de plataforma de união entre os dois elementos anteriormente referidos.

A câmara escolhida para aquisição das imagens foi uma *action cam* (GoPro Hero3 Silver), produzida pela *GoPro, Inc*, empresa Americana que desenvolve, produz e comercializa câmaras pessoais de alta-definição, normalmente utilizadas em fotografia e vídeo de desportos extremos. Estas câmaras são conhecidas por serem compactas, leves, robustas e com variadas possibilidades de acopulação/montagem, tais como em quadcopters, aviões, carros, barcos, etc. Estas câmaras têm a capacidade de capturar fotografias paradas e/ou vídeo alta-definição, através de uma lente grande angular tendo a capacidade de serem configuradas para trabalharem automaticamente com o mínimo de intervenção e/ou remotamente.

A versão Hero 3 Silver tem a vantagem de ter equipado no seu hardware um sistema *Wi-Fi*, tendo em modo de fotografia que de acordo com a referência [10], as seguintes especificações:

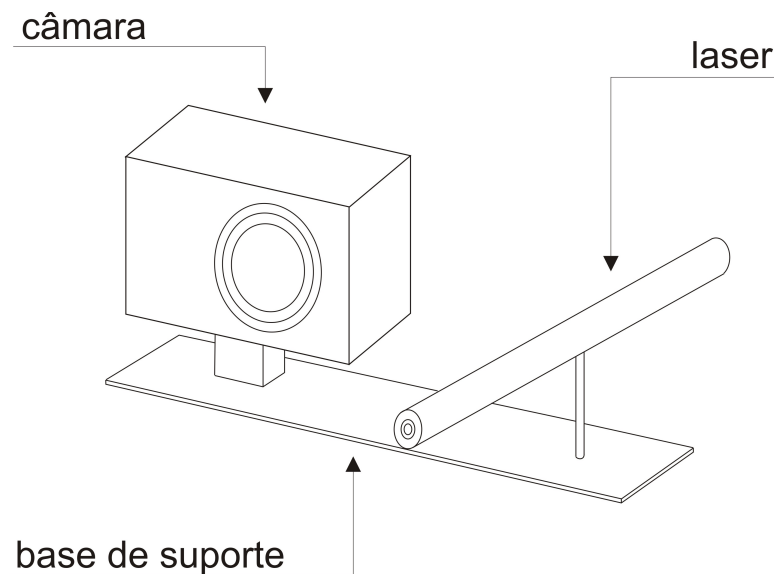


Tabela 3.2: Especificações da câmara GoPro em modo de Fotografia.

Resolução	Campo de visão (FOV)	Resolução da tela
10 MP	Ultralargo	3680 x 2760
7 MP	Ultralargo	3072 x 2304
5 MP	Médio	2624 x 1968

Sendo também capaz de produzir fotos em Sequência (10 fotos/s ; 5 fotos/s ; 3 fotos/s) ou em Time-Lapse, modo em que tira fotografias de acordo com um determinado intervalo de tempo distanciado de 0,5 , 1, 2, 5, 10, 30 ou 60 segundos. Pela versatilidade oferecida por esta câmara e pelo facto de ser possível o controlo remoto do sistema, foi decidida a utilização desta câmara como sensor para a captação das imagens.

Na figura 3.2 está apresentado um *concept* do sensor proposto que, tal como referido anteriormente, será constituído por uma câmara, um emissor de laser e uma base que une os elementos do sensor. Este *concept* fornece, ao sistema de aquisição tridimensional proposto, todos os elementos de interacção com o cenário em estudo necessários para que seja possível a aquisição dos dados necessários na utilização da tecnologia de aquisição proposta, triângulação.

Figura 3.2: *Concept* do sensor a utilizar no processo.

### 3.3 Software

A componente de software desta proposta é constituída por todos os componentes que contribuem para a comunicação, percepção e cálculo dos elementos capturados por parte do *hardware*. Observando a figura 3.1 é possível concluir que a componente software do sistema proposto é dividido em 3 sub-sistemas: Sistema de Comunicação, Sistema de Visão Computacional e Sistema de Cálculo. Nesta secção ir-se-á propôr quais as funções e funcionamento de cada um destes sub-sistemas.

#### 3.3.1 Sistema de Comunicação

Este sistema é composto pelos elementos que permitem a comunicação entre o sensor e o computador. Este sistema é a base da aquisição à distância, tornando possível a interacção com o sensor via *Wi-Fi*.

A componente de *Hardware* escolhida, nomeadamente a câmara *GoPro* permite que o sensor tenha uma componente *Wi-Fi* que, quando ligada, tem a capacidade de criação de uma rede *ad hoc*. Com a criação desta rede é possível que dispositivos com sistemas *Wi-Fi* se liguem a esta rede e, por sua vez, ao sensor em questão. A *GoPro, Inc.* fornece gratuitamente aos seus utilizadores uma aplicação para telemóvel que utiliza o sistema *Wi-Fi* integrado da câmara, para possibilitar não só o total controlo das funcionalidades da mesma, como também o acesso às fotografias e vídeos por ela captados [10]. Utilizando engenharia inversa é possível criar uma reprodução dos comandos utilizados pela aplicação de telemóvel fornecida pela *GoPro, Inc.*, de forma a que seja possível comandar o sensor, por exemplo, através de um computador que se conecte à rede *ad hoc*.

Após a realização de alguma pesquisa em fóruns de utilizadores destas câmaras [11, 12, 7] foi possível encontrar o formato dos comandos enviados para a rede de forma a que a câmara saiba que função desempenhar.

É possível enviar os comandos para a câmara sob o formato de *Uniform Resource Locator* (URL) a seguir apresentado:

`http://[CâmaraIP]/[cameraorbacpac]/[COMANDO]?t=[PASSWORD]&p=%[NUMERO]`

Entre os elementos que constituem o URL, existem alguns que se mantêm constantes de comando para comando: o endereço do *Internet Protocol* (IP) da câmara é estático e por predefinição tem o valor de "10.5.5.9:8080" e a *PASSWORD* é a palavra passe da rede *ad hoc* criada pela câmara. Em seguida, na tabela 3.3 estão apresentados alguns dos comandos usados para controlar a câmara.

Tabela 3.3: Alguns comandos usados para controlar a câmara *GoPro* via *WiFi*

TIPO	COMANDO	NÚMERO	FUNÇÃO
bacpac	PW	00	Desligar GoPro
bacpac	PW	01	Ligar GoPro
bacpac	SH	01	Activar Obturador
bacpac	SH	00	Parar
camera	CM	00	Modo Video
camera	CM	01	Modo Fotografia
camera	CM	02	Modo Disparo Rápido
camera	CM	00	Modo <i>TimeLapse</i>
camera	PR	00	Resolução 11mP Largo
camera	PR	01	Resolução 8mP Médio
camera	PR	02	Resolução 5mP Largo
camera	PR	03	Resolução 5mP Médio
camera	TI	00	Intervalo <i>TimeLapse</i> 0,5 seg.
camera	TI	01	Intervalo <i>TimeLapse</i> 1 seg.
camera	TI	05	Intervalo <i>TimeLapse</i> 5 seg.
camera	TI	0a	Intervalo <i>TimeLapse</i> 10 seg.
camera	TI	1e	Intervalo <i>TimeLapse</i> 30 seg.
camera	TI	3c	Intervalo <i>TimeLapse</i> 60 seg.

Os comandos referidos na tabela 3.3 serão utilizadas pelo *Sistema de Comunicação* de forma a controlar as acções tomadas pela câmara do sensor. Como a câmara vai guardando em memória as fotografias da aquisição do cenário é necessário aceder à localização das mesmas via *Wi-Fi*. Esse acesso, é executado acedendo, através do IP da câmara, à pasta na qual esta vai armazenando as fotografias. Na figura 3.3 é possível observar o *webserver* criado pela câmara *GoPro* de forma a que seja possível aceder-se, através da aplicação, aos ficheiros por ela criados.



Figura 3.3: Acesso através de um *browser* ao *webserver* criado pela câmara *Gopro* que permite o acesso aos ficheiros por ela criados.

No caso do *Sistema de Comunicação*, este vai aceder directamente à pasta do *webserver* onde estão

### 3. PROPOSTA

guardadas as fotografias e, a partir desse ponto, executará o *Download* das fotografias do cenário em estudo, uma a uma.

#### 3.3.2 Sistema de Visão Computacional

O Sistema de Visão Computacional é responsável por tornar as imagens recebidas pelo *Sistema de Comunicação* e transformá-las, de forma a que estas possam ser interpretadas pelo *Sistema de Cálculo*. Cabe a este sistema localizar o laser na imagem e devolver, sob a forma de coordenadas dos pixels da imagem, as posições às quais o laser incide sobre a superfície do cenário em estudo. Na figura 3.4 está apresentado o esquema geral do funcionamento deste sistema.

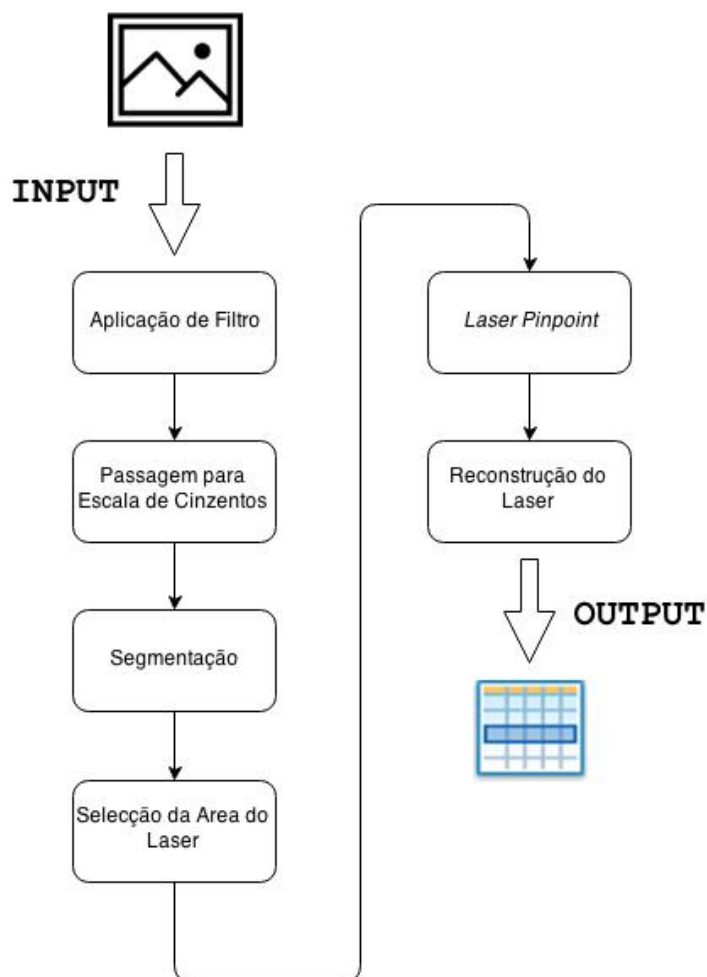


Figura 3.4: Esquema geral do funcionamento do Sistema de Visão Computacional.

De seguida estão detalhados cada um dos passos efectuados na avaliação de uma imagem por parte do *Sistema de Visão Computacional*, onde são justificadas as escolhas e os algoritmos propostos para a implementação do sistema proposto.

### 3.3.2.1 Aplicação de Filtro

Tendo em consideração o estudo realizado no estado da arte em Processamento de Imagem (capítulo 2.2), podemos concluir que as imagens obtidas pela câmara conterão ruído. A quantidade de ruído existente na imagem irá interferir directamente a capacidade de decisão do *Sistema de Visão Computacional* pela existência de pixéis com valores que não reflectem a verdadeira intensidade existente no cenário real. É por isso fulcral, para que o *Sistema de Visão Computacional* ofereça um desempenho desejado que a quantidade de ruído existente na imagem seja atenuado/eliminado.

Para que se possa proceder à escolha de um filtro que melhore a percepção dos dados da imagem é necessário primeiro perceber-se: qual o tipo de ruído que mais provavelmente se encontrará e quais as características da imagem que se esperam obter.

Tendo em consideração que, durante as aquisições, para se conseguir melhores resultados é necessário que o ambiente do cenário seja escuro de forma a que o laser se destaque em termos de intensidade em relação ao resto do cenário, o tipo de ruído que se espera encontrar em maior quantidade será o ruído *salt-n-pepper* (ruído impulsivo), sendo portanto esperado encontrar-se na imagem pixeis que se destacam por serem mais claros que a sua vizinhança. Outro factor muito importante na escolha do filtro a utilizar, é a necessidade de preservar os contornos da imagem, pois o laser projectado na imagem será o mais fino possível e apresentará contornos salientes.

Sabendo que os filtros lineares suavizam toda a imagem, ao desfocarem os contornos da imagem deixam de ser uma opção válida para o objectivo pretendido. Por sua vez, os filtros não-lineares, nomeadamente o filtro de mediana, apresenta um grande desempenho na eliminação de ruído *salt-n-pepper* e têm a capacidade de suavizar a imagem. Por esta razão será escolhido a utilização de um filtro de mediana como forma de reduzir/eliminar o ruído presente nas imagens capturadas pelo sensor.

### 3.3.2.2 Passagem para Escala de Cinzentos

Esta fase do *Sistema de Visão Computacional* é uma fase preparatória da *Segmentação* e consiste na transformação da imagem já filtrada numa imagem em escala de cinzentos.

Sabendo que cada pixel da imagem é constituído por três canais (figura 3.5a), *Red Green and Blue* (vermelho, verde e azul) (RGB), e que a soma dos 3 canais forma a cor do pixel que observamos é possível criar tons de cinzento ao introduzirmos em cada canal o mesmo valor numérico, sendo (0,0,0) o preto total e (255,255,255) o branco.



(a) Representação da estrutura de um pixel.



(b) Escala de cinzentos.

Figura 3.5: Esquema explicativo da estrutura de um pixel (a) e da representação de uma escala de cinzentos (b).

O cálculo do valor de um pixel em escala de cinzentos pode ser obtido através da equação 3.1

$$R = \frac{P_R + P_G + P_B}{3} \quad (3.1)$$

onde  $P_R$  é o valor do canal Vermelho do pixel,  $P_G$  é o valor do canal Verde e  $P_B$  é o valor do canal Azul.

Como o feixe laser emitido terá cor verde, em vez de se fazer uma média do valor dos 3 canais e atribuir o resultado dessa média a cada um dos canais para obter o grau de cinzento correspondente, considerar-se-á apenas o valor do canal verde que será aplicado nos três canais, obtendo assim uma escala de cinzentos que avalia o quão verde o pixel em avaliação é. Assim, quanto maior for o valor do canal verde do pixel avaliado, mais claro ele se tornará na imagem em escala de cinzentos, tal como é possível observar na figura 3.5b.

### 3.3.2.3 Segmentação

Após se ter transformado a imagem em escala de cinzentos, é necessário proceder à segmentação da imagem, de forma a que esta seja simplificada e o *Sistema de Visão Computacional* possa determinar quais os pixels que poderão pertencer à passagem do laser e quais os que não irão corresponder a essa passagem. Para isso é executada uma binarização da imagem, de forma a que os pixels apresentem um de dois valores, 0 ou 255. Através do método de *Thresholding* referido na secção 2.3 é possível executar a binarização da imagem, fazendo com que todos os pixels que apresentem valores inferiores

ou iguais ao valor de *threshold* fiquem com valor 0, ou seja, preto, e os pixels com valores superiores ao valor de *threshold* fiquem com valor 255 (branco).

Desta maneira, o algoritmo de binarização percorre a imagem pixel a pixel atribuindo um novo valor ao pixel consoante este esteja abaixo ou acima do valor de *threshold* (175). A necessidade da imagem estar em escala de cinzentos antes de ser aplicada a segmentação vem de só assim ser possível determinar se o novo valor do pixel será preto ou branco por observação de um dos canais do pixel.



Figura 3.6: Exemplo de utilização da binarização numa imagem.

Na imagem 3.6 é possível observar um caso de binarização de uma imagem e é fácil perceber as vantagens que esta traz ao Sistema de Visão Computacional, facilitando a percepção de objectos na imagem, no caso do exemplo a ficha de alimentação de um portátil. No caso do laser, o objectivo da binarização será a obtenção de uma linha/curva ao longo da imagem, de forma a facilitar a localização do laser na imagem.

#### 3.3.2.4 Selecção da Área do Laser

Com a imagem binarizada e realizada a inversão das cores da imagem, é agora possível ao *Sistema de Visão Computacional* determinar a posição de localização do laser na imagem.

Nesta fase, o algoritmo de *Selecção da Área do Laser* determinará, por avaliação da imagem, a largura ocupada por toda a secção atravessada pelo laser na imagem, criando desta maneira uma nova imagem com altura igual à original, mas de largura igual ao espaço horizontal que o feixe laser ocupa ao longo de toda a sua passagem na imagem original.

De forma a determinar essa área, o algoritmo de *Selecção da Área do Laser* começa por executar uma contagem do número de pixels pretos existente em cada coluna da imagem, criando um histograma semelhante ao apresentado na figura 3.7.

### 3. PROPOSTA

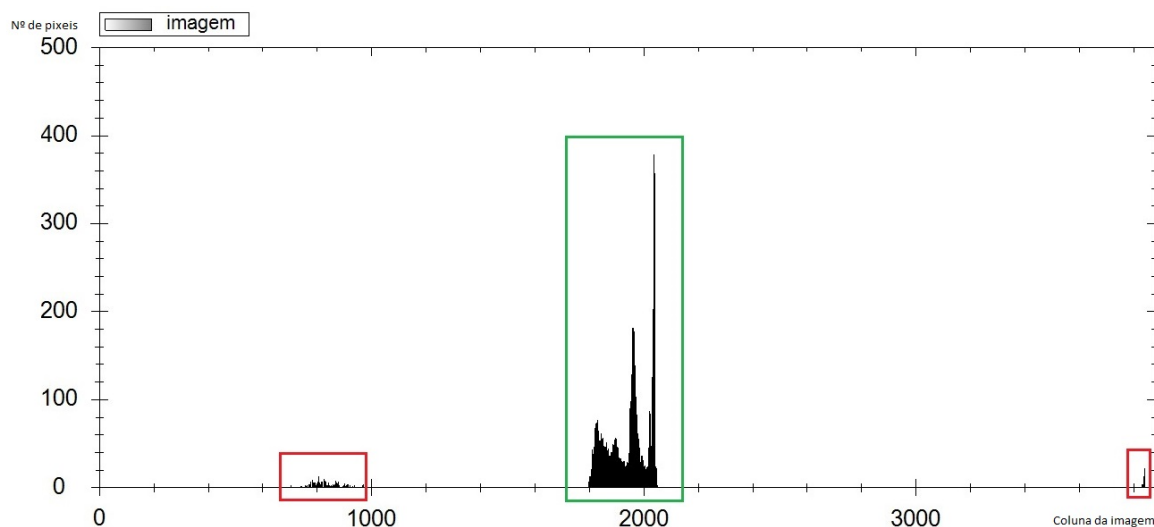


Figura 3.7: Histograma de contagem de ocorrência de pixels pretos por coluna da imagem.

Como é possível verificar existem 3 agrupamentos visíveis no histograma, sendo que o maior agrupamento pertencerá à área ocupada pelo laser na imagem. Os agrupamentos mais pequenos são zonas de ruído que não foram eliminadas pelo processo de filtragem. A razão para se saber que o maior agrupamento de pixels pretos pertence à área que o laser ocupa na imagem, deve-se a se saber que este ocupa a imagem em toda a sua altura, sendo portanto natural que haja uma maior quantidade de pixels pretos nas colunas da imagem passadas pelo laser.

Após ter sido obtido o histograma, o algoritmo realiza uma contagem dos pixels pertencentes a cada agrupamento e selecciona o agrupamento que apresentar o valor máximo de pixels pretos. Com essa selecção, o algoritmo ganha a capacidade de definir os limites do agrupamento de forma a saber qual a coluna onde primeiro é detectada a presença do laser e a coluna onde é detectada a sua última presença, ficando assim definida a área de passagem do laser na imagem.

#### 3.3.2.5 *Laser Pinpoint*

Após já termos criada a nova imagem com a largura do laser, é-nos agora possível realizar o *pinpoint* da localização do laser ao longo de toda a imagem, de forma a se obter uma representação do laser como se ele apenas tivesse um pixel de largura.

Para que isso seja possível, este algoritmo terá de percorrer a nova imagem linha a linha de forma a determinar qual o pixel que se considerará como pixel de passagem do laser.

Em cada linha o algoritmo irá somar todas as posições que contenham valor 0 (pixels pretos) e no final, antes de passar para a próxima linha, calcula a média desses pontos, obtendo assim o valor médio dos pontos que este considera que contêm passagem do laser (pixels pretos).

Na figura 3.8 está representado o fluxograma representativo do funcionamento do algoritmo de *pinpoint*.



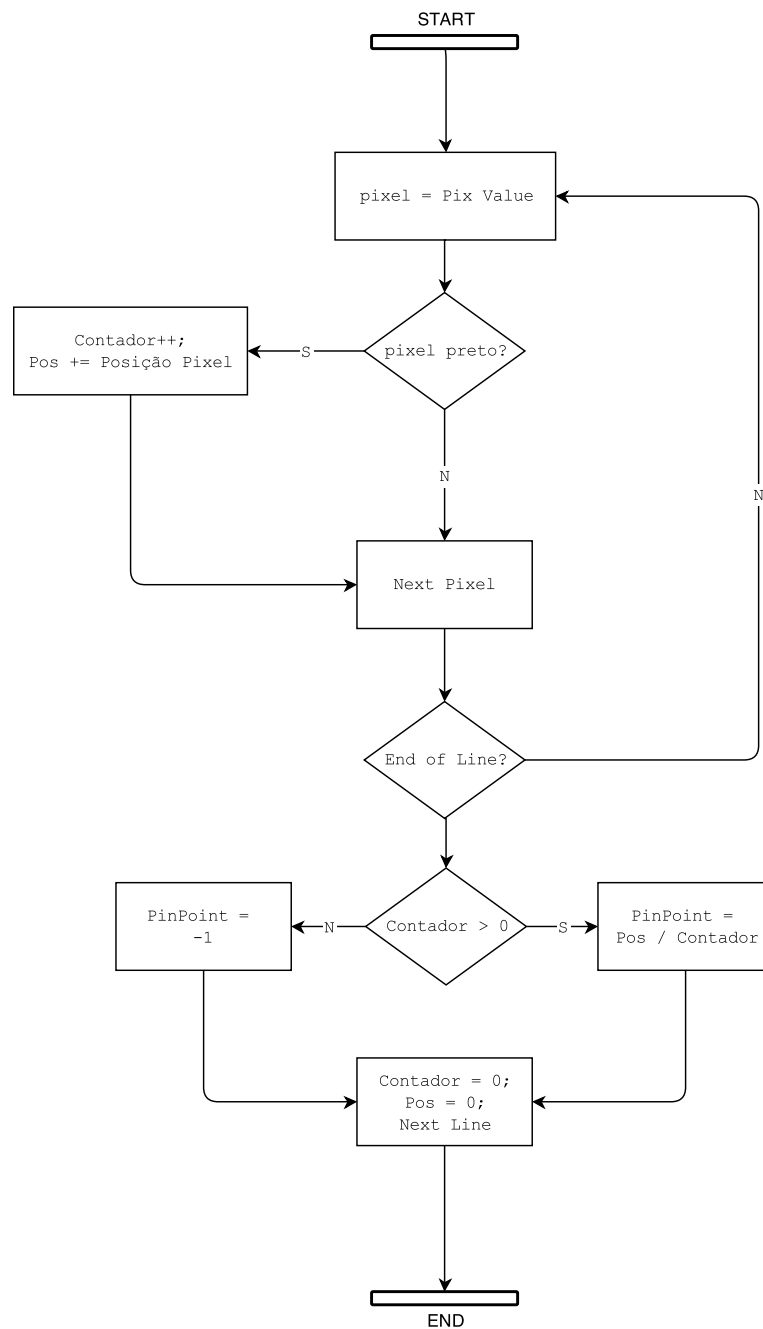


Figura 3.8: Fluxograma representativo da utilização do algoritmo de *pinpoint*.

Este algoritmo começa por guardar numa variável(*pixel*) o valor do pixel actual, caso o pixel seja preto, um contador é incrementado e soma-se o valor da posição do pixel preto actual a uma variável (*Pos*), variável essa que contém a soma de todos os pixels pretos da linha em estudo, até ao pixel actual. Depois é incrementada a posição do apontador, de forma a apontar para o próximo pixel da linha. No caso do pixel referido no passo anterior não ser preto, o contador e a variável *Pos* não são mexidas, apenas se incrementando a posição do apontador. No caso de a nova posição não ser *End of Line*

### 3. PROPOSTA

(fim de linha) o algoritmo entra num ciclo que realiza a verificação dos pixels pretos da linha, até que este seja encontrado. Quando for encontrado, o *End of Line*, caso o contador tenha valor igual a zero, ou seja, não existirem pixels pretos na linha, o valor guardado será *NaN(Not-a-Number)*, no caso do contador ser maior que zero o valor guardado será a divisão da soma de todos as posições dos pixels pretos *Pos* pelo número de pixels pretos. O valor resultante será o valor definido como a posição exacta de passagem do laser. Depois de definida essa posição, as variáveis são reiniciadas e o apontador passa a apontar para a próxima linha da imagem e o ciclo repete-se até não haver linha seguinte. Na figura reffig:trans está representado um exemplo de *laser pinpoint*.



Figura 3.9: Exemplo de aplicação de *laser pinpoint*.

#### 3.3.2.6 Reconstrução do Laser

Para que seja possível a reconstrução do laser é necessário seguir os passos apresentados na figura 3.10 abaixo apresentada.

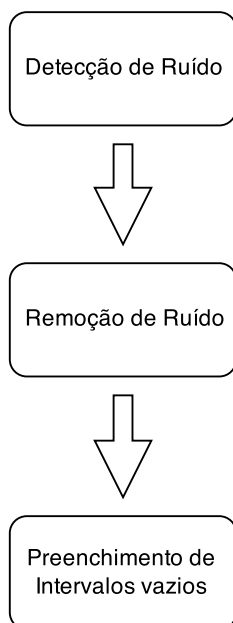


Figura 3.10: Paços utilizados para a reconstrução do Laser.

Como foi possível verificar, na fase de *pinpoint* do laser (secção 3.3.2.5), é possível que existam zonas do *array* do algoritmo que, em vez de terem a posição de onde o laser deveria passar, têm *NaN*. Outra situação que pode ocorrer, aquando da existência de ruído na imagem, é a perturbação da precisão do algoritmo de *pinpoint*, criando ruído de informação sob a forma de flutuações ou desvios no posicionamento correcto do valor do pixel de passagem do laser. Para ultrapassar estes problemas, na fase de *Reconstrução do Laser*, e tal como apresentado na figura 3.10 são executadas 3 tarefas de forma a tentar recuperar de uma forma mais precisa as coordenadas dos pixels de passagem dos laser na imagem.

A primeira tarefa consiste na *Deteccção de Ruído*. Para se poder distinguir entre zonas de ruído, já que não existe nenhuma forma correcta ou errada por parte da linha laser, é necessário determinar o que queremos obter e o que esperamos obter. Sabe-se que o laser é suposto apresentar-se na imagem como uma linha/curva contínua ao longo das superfícies do cenário. É por isso esperado ver-se os pixels seleccionados variarem pouco em relação aos seus vizinhos. Na imagem 3.11 é possível verificar uma comparação entre uma linha com os valores dentro do esperado (figura 3.11a) e outra com valores que poderão ser variações causadas por ruído (figura 3.11b).

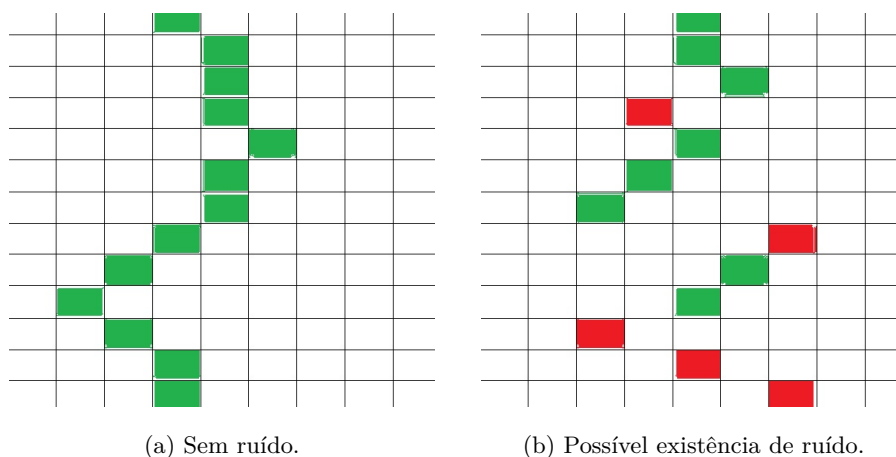


Figura 3.11: Comparação entre imagem de acordo com o padrão de laser esperado (a) e com possível existência de ruído (b).

Apesar de o princípio de detecção de ruído ser o que se apresenta no exemplo da figura 3.11b, este não é suficiente para determinar a existência ou não existência de ruído. Foi por isso necessário determinar um conjunto de regras para, baseado no princípio anteriormente referido, encontrar intervalos de ruído no *pinpoint* do laser.

Essas regras podem ser observadas no fluxograma apresentado na figura 3.12 que representa a avaliação do *array* criado pelo algoritmo de *pinpoint* e a determinação de quais os intervalos que devem ser considerados como ruído e quais os que devem ser determinados como flutuações normais dos valores. A detecção de ruído é então determinada pela existência de um intervalo de pixels de ruído maior que 10. Para se determinar o intervalo de ruído é necessário que o seu início e fim sejam determinados,

### 3. PROPOSTA

por isso é fulcral que se volte a encontrar um conjunto de 5 pixels que não sejam considerados como possível ruído. O facto de ser necessário encontrar esse conjunto de pixels "normais" e não apenas 1 desses pixels deve-se à necessidade de garantir que não se continua no intervalo de ruído, pois é normal ocorrerem zonas de ruído constituídas por pequenos conjuntos de pixels considerados "normais".

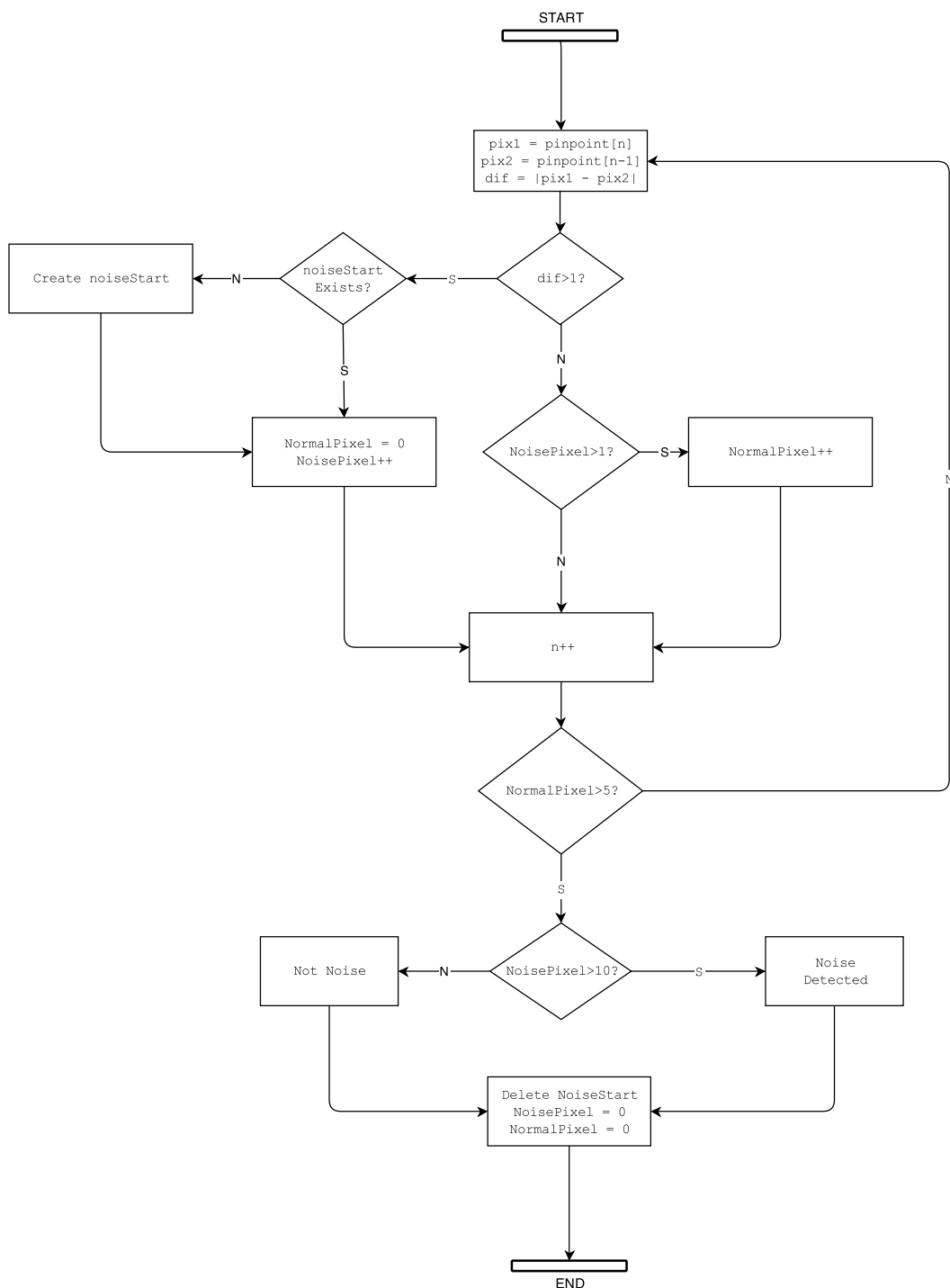


Figura 3.12: Fluxograma representativo das regras de determinação de intervalos de ruído.

Tendo sido determinados os intervalos de ruído no *array* criado pelo método de *pinpoint*, é executado o processo de eliminação desses intervalos, passando a apresentar um valor *NaN*, criando desta forma buracos nas zonas onde antes existia ruído. A esta fase é chamada a *Remoção de Ruído* que está apresentana na figura 3.10.

A fase anterior é uma intervenção destrutiva dos dados obtidos, elimina os dados considerados errados. Assim, neste momento, temos um laser com buracos causados pela eliminação de ruído e buracos existentes pela insuficiência de dados existente. Desta forma, e tal como o nome indica, nesta fase é executado o *Preenchimento de Intervalos Vazios* de forma a termos uma linha/curva laser completa. O conceito por trás deste algoritmo é simples e consiste em ligar os espaços vazios preenchendo os mesmos com valores que progressivamente tenderão para o valor final do intervalo.

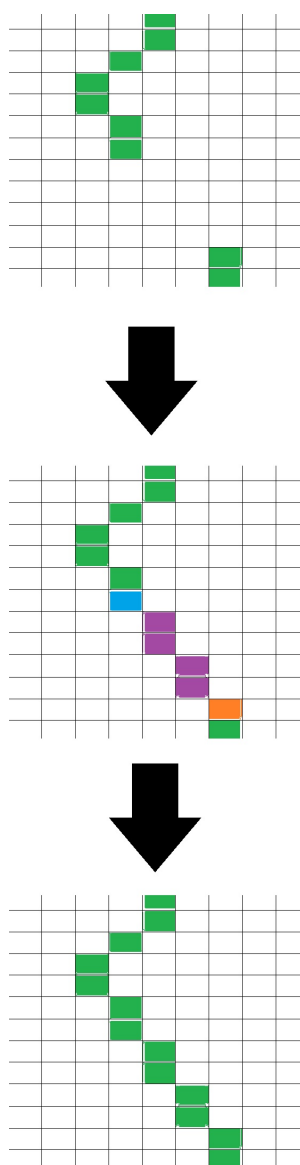


Figura 3.13: Exemplo de aplicação do processo de preenchimento de espaços vazios.

### 3. PROPOSTA

Na imagem 3.13 é possível observar um exemplo do funcionamento desta fase. Num primeiro estado verifica-se a existência de um espaço vazio, no segundo estado, a azul e a laranja são identificados respectivamente os pontos de início e fim do espaço vazio. É então preenchido o intervalo de forma a que os pixels se encontrarem entre o intervalo traçado pelo início e o fim (roxo) fiquem ligados. No último estado verifica-se o resultado final do preenchimento do espaço vazio.

#### 3.3.3 Sistema de Cálculo

O *Sistema de Visão Computacional* entrega ao *Sistema de Cálculo* uma tabela com a informação das várias aquisições realizadas durante o estudo do cenário. A tabela recebida segue uma regra de formato de forma a que o *Sistema de Cálculo* saiba como extrair a informação da tabela.

O formato utilizado nas tabelas segue os seguintes regras:

**Colunas** representam os valores dos pixels em que foi detectada a passagem do laser na imagem ao longo da sua altura.

**Linhas** representam os vários estudos, ou as várias imagens analisadas pelo *Sistema de Visão Computacional* no estudo do cenário.

Desta forma é normal que existam na tabela 2880 posições, que correspondem à altura das imagens que serão utilizadas, sendo que a primeira coluna corresponde ao pixel mais alto do cenário e a última coluna correspondente ao mais baixo. Desta forma é possível reproduzir as coordenadas (X,Y) do pixel que se quer avaliar, de forma a calcular a distância nesse ponto, sendo X correspondente ao valor na célula da coluna e Y o número da coluna.

Ao se realizarem  $n$  aquisições durante o estudo de um cenário, ir-se-á obter  $n$  linhas na tabela, cada uma dessas linhas com um conjunto de 2880 pontos associados, tal como referido no parágrafo anterior. Ao se realizarem as aquisições com um intervalo rotação constante, é possível determinar qual o ângulo horizontal de cada linha da tabela através de uma simples multiplicação.

## 3.4 Sinopse

Neste capítulo, foram escolhidos quais os métodos e tecnologias que integraram o processo de forma a cumprir os requisitos previamente traçados. É referido também como se pretende que o processo se comporte nomeadamente a relação entre as componentes *Hardware* e *Software* do processo. No capítulo seguinte, vai-se realizar a aplicação do processo proposto, de forma a poder-se proceder à validação do mesmo, mediante a criação de um protótipo de um sistema de percepção tridimensional.

# CAPÍTULO 4

## Validação

---

<b>4.1</b>	<b>Método de Validação</b>	<b>41</b>
<b>4.2</b>	<b>Protótipo</b>	<b>42</b>
4.2.1	Cálculo das distâncias	44
4.2.2	Cálculo do Ângulo da Câmara	47
<b>4.3</b>	<b>Experimentação</b>	<b>49</b>
4.3.1	Teste de Calibração	50
4.3.2	Teste de percepção de profundidade de uma linha laser	51
4.3.3	Teste de percepção tridimensional por rotação do sensor	53
<b>4.4</b>	<b>Sinopse</b>	<b>54</b>

---

### 4.1 Método de Validação

De forma a se poder validar o modelo proposto no capítulo anterior, foi criado um protótipo que será experimentado em cenários de espaços interiores com iluminação artificial controlada. Essas experiências foram executadas de forma a testar a capacidade de percepção espacial do modelo e a sua performance, nomeadamente a precisão do modelo, suas limitações e por fim comprovar a possibilidade do uso deste tipo de modelos como *scanner* tridimensional desses espaços.

Foram utilizados vários cenários para o estudo de calibração do sensor, de forma a relizar afinações nos cálculos e proceder à detecção de possíveis obstáculos à percepção tridimensional e à criação de *point-clouds* tal como é desejado. Após esses testes de calibração foram estudados dois cenários com diferentes objectivos de aquisição e de validação.

No primeiro cenário teve-se o objectivo de testar apenas a capacidade de análise de profundidade com o protótipo estático, avaliando nesse sentido o espaço de uma forma bidimensional, sendo avaliadas apenas a profundidade e altura do cenário proposto. O cenário que se estudou consistiu numa parede

#### 4. VALIDAÇÃO

com uma viga no topo e uma cama encostada à parede.

O segundo cenário estudado foi uma garagem. Aqui já se procedeu à análise de todo o volume da mesma, sendo introduzida uma terceira variável, uma rotação em torno do eixo dos YY (atitude). Rodando o protótipo em torno de si próprio quis-se obter uma *point-cloud* de todo o cenário em redor da câmara e não apenas uma linha do cenário, como foi testado no primeiro cenário.

### 4.2 Protótipo

O protótipo construído, com vista no cumprimento dos objectivos traçados na secção 1.3 e baseado na proposta definida no capítulo 3, é constituído por uma componente sensorial (*Hardware*) e outra perceptual(*Software*). A primeira consiste num dispositivo que capta o cenário em estudo, podendo ser observado na figura 4.1.

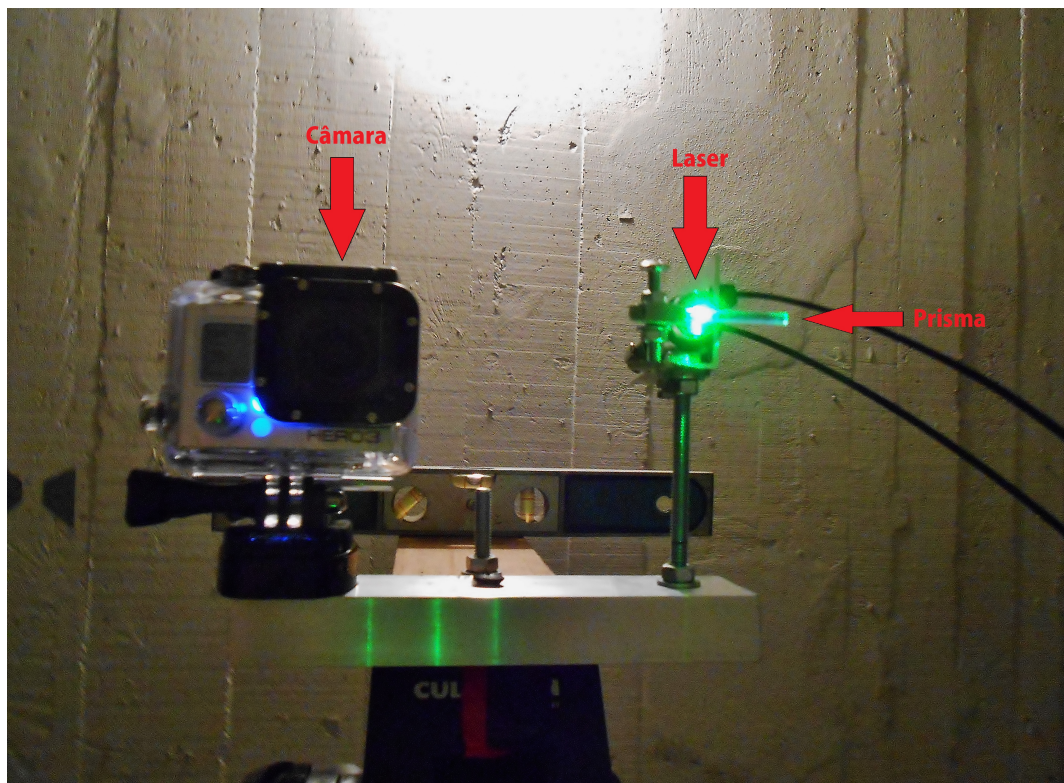


Figura 4.1: Sensor usado para a captação dos cenários e seus elementos constituintes.

Na mesma figura é possível observar os elementos constituintes do sensor, nomeadamente a câmara (GoProHero 3 Silver), um emissor luz laser e um prisma. A câmara tem a função de captar as imagens do cenário, mais precisamente, tem o objectivo de ir memorizando sob a forma de imagens, os pontos de incidência da linha laser nas superfícies do cenário em estudo. Os dois últimos elementos do sensor, emissor laser e o prisma, formam um sistema que trabalha em conjunto, de forma a ser projectado no



cenário a linha laser vertical desejada.

Tal como pode ser observado no esquema da figura 4.2, o prisma tem a capacidade de refratar a luz por ele atravessada, dispersando assim a luz emitida pelo emissor laser e formando uma linha ao longo de uma parede vertical.

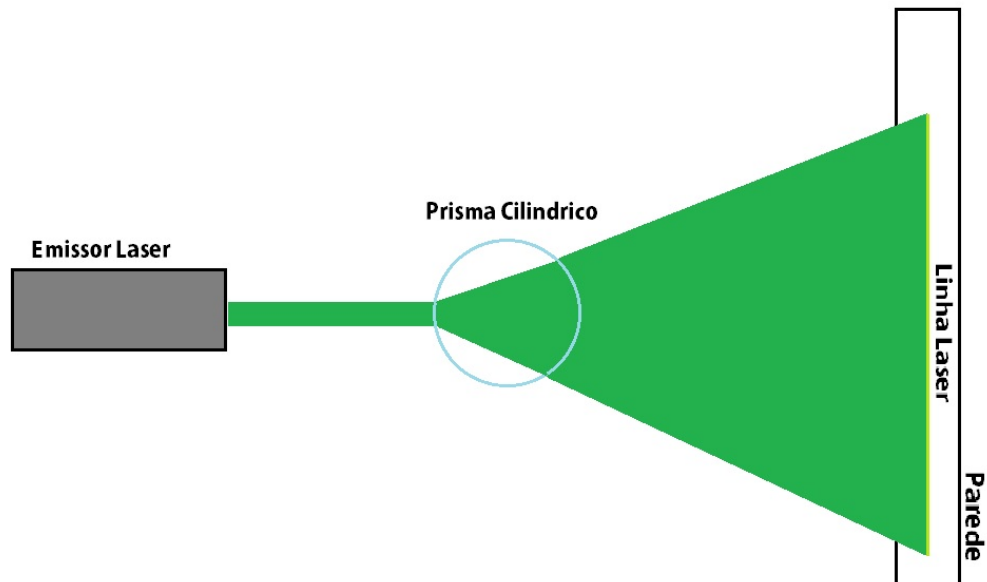


Figura 4.2: Esquema representativo da refração da luz laser através da utilização de um prisma cilíndrico.

A relação de distâncias existente entre os vários elementos do sensor está represenada na figura 4.3. Estas distâncias foram obtidas na fase de calibração e são utilizadas no cálculo de distâncias como constantes.

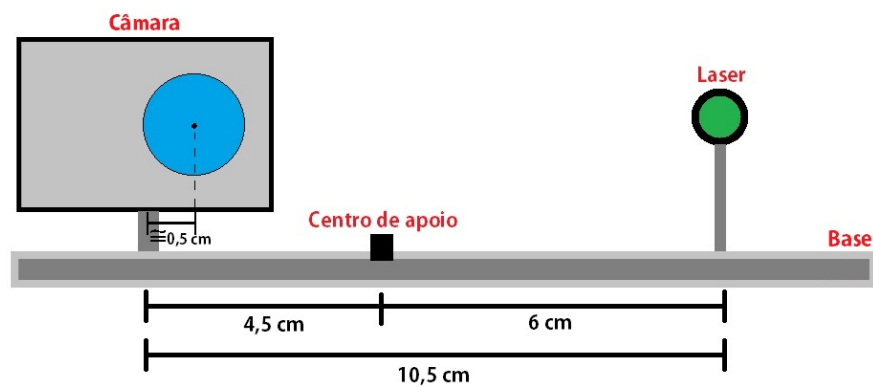


Figura 4.3: Esquema representativo da relação de distâncias existente entre os elementos do sensor.

#### 4. VALIDAÇÃO

O *centro de apoio* representado na figura 4.3 é o ponto no qual a base do sensor é apoiada durante as experimentações sendo também o seu centro de rotação.

##### 4.2.1 Cálculo das distâncias

O protótipo, de forma a calcular a profundidade de um cenário, utiliza 3 formas distintas de cálculo com base em triangulação, de forma ser capaz de perceber as distâncias.

A primeira e mais simples baseia-se no conceito de triangulação, no teorema de pitágoras e representa a distância de referência, situação em que o ponto do laser em estudo se situa no centro geométrico da imagem.

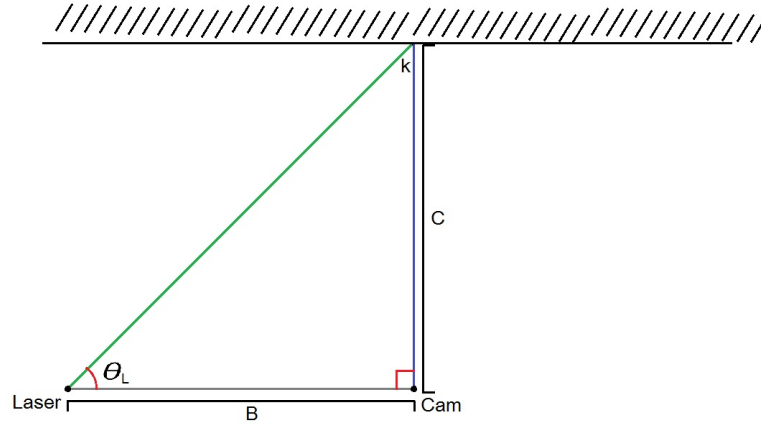


Figura 4.4: Esquema de cálculo da distância de referência.

Na figura 4.4 é possível calcular a distância ao ponto  $k$  do objecto em que o laser está a incidir utilizando:

$$C = B \times \cos(\theta_L) \quad (4.1)$$

Nesta equação, sabe-se *a priori*, através da calibração os valores de  $B$  (distância entre o laser e a câmara) e o valor de  $\theta_L$ , que é o ângulo que o laser faz com a base. Com a equação 4.1 conseguiu-se calcular a distância de referência, agora para situações em que o objecto está mais próximo do que a distância de referência em relação à câmara, é necessário proceder a um tipo de cálculo diferente, pois, ao realizarmos a triangulação entre as posições do laser, da câmara e do ponto de incidência do laser no objecto, já não obtemos um triângulo rectângulo mas sim uma representação do tipo da apresentada na figura 4.5.

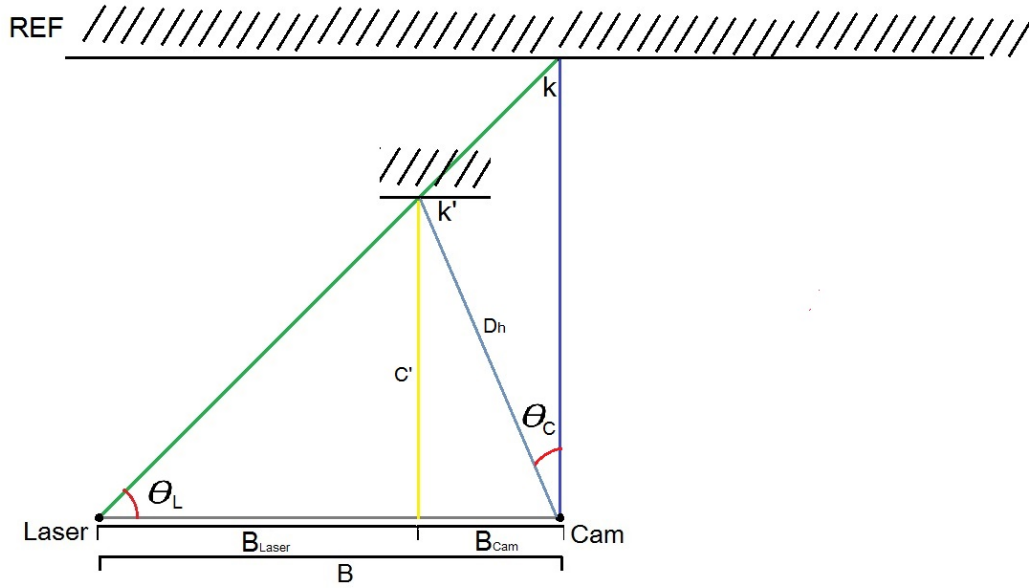


Figura 4.5: Esquema de cálculo para distâncias inferiores à distância de referência.

$$\left\{ \begin{array}{l} B_{Laser} = B - B_{cam} \\ \tan(\Theta_L) = \frac{C'}{B_{Laser}} \\ \cos(90^\circ - \Theta_{Cam}) = \frac{B_{cam}}{D_h} \\ \tan(90^\circ - \Theta_{Cam}) = \frac{C'}{B_{cam}} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} B_{Laser} = B - B_{cam} \\ C' = \tan(\Theta_L) \times (B - B_{cam}) \\ D_h = \frac{B_{cam}}{\cos(90^\circ - \Theta_{Cam})} \\ \tan(90^\circ - \Theta_{Cam}) = \frac{\tan(\Theta_L) \times (B - B_{cam})}{B_{cam}} \end{array} \right. \quad (4.2)$$

Pegando no último ramo do sistema de equações 4.2 e resolvendo-o em ordem a  $B_{cam}$  obtemos:

$$B_{cam} = \frac{B \tan(\Theta_L)}{\tan(90^\circ - \Theta_{Cam}) + \tan(\Theta_L)} \quad (4.3)$$

Ao substituir a equação 4.3 no segundo ramo do sistema de equações 4.2 de forma a obter  $C'$  em função dos parâmetros conhecidos da calibração  $\Theta_L$ ,  $B$  e do parâmetro obtido por avaliação da imagem  $\Theta_{Cam}$  obtemos:

$$C' = B \tan(\Theta_L) - \frac{B \tan(\Theta_L)^2}{\tan(90^\circ - \Theta_{Cam}) + \tan(\Theta_L)} \quad (4.4)$$

Obtém-se assim o sistema de equação 4.5 que representa a solução dos parâmetros desconhecidos  $C'$  e  $B_{cam}$  em função dos parâmetros conhecidos, de forma a ser possível o cálculo da distância,  $D_h$ , entre a

#### 4. VALIDAÇÃO

câmara,  $Cam$ , e o ponto de incidência do laser no objecto,  $k'$ :

$$\begin{cases} C' = B \tan(\Theta_L) - \frac{B \tan(\Theta_L)^2}{\tan(90^\circ - \Theta_{Cam}) + \tan(\Theta_L)} \\ B_{cam} = \frac{B \tan(\Theta_L)}{\tan(90^\circ - \Theta_{Cam}) + \tan(\Theta_L)} \\ D_h = \frac{B_{cam}}{\cos(90^\circ - \Theta_{Cam})} = \frac{C'}{\sin(90^\circ - \Theta_{Cam})} = \sqrt{(C')^2 + (B_{cam})^2} \end{cases} \quad (4.5)$$

Para os casos em que a distância entre a câmara e o ponto de incidência do laser se encontrem a uma distância superior à distância de referência obteremos um esquema de cálculo como o representado na figura 4.6.

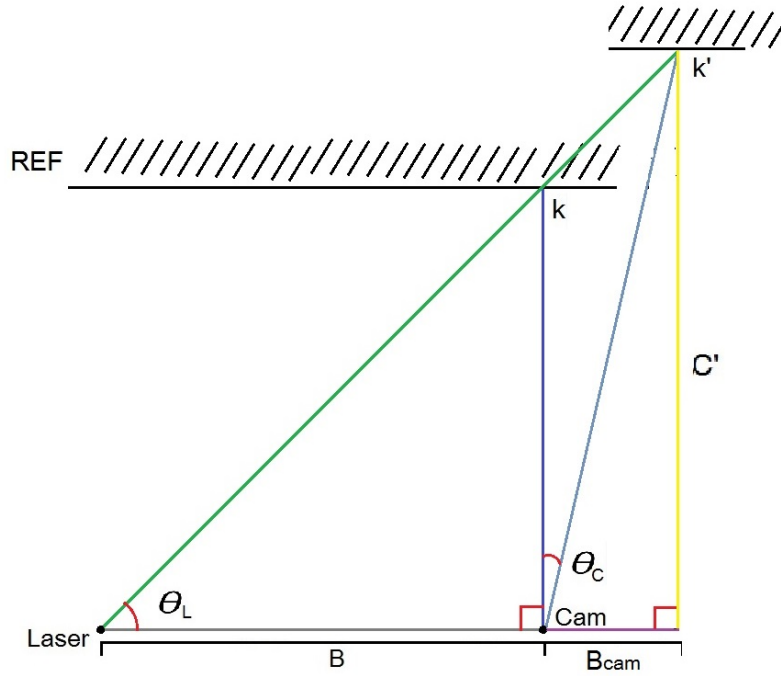


Figura 4.6: Esquema de cálculo para distâncias superiores à distância de referência.

Nesta situação, o método de cálculo da distância  $D_h$ , é semelhante ao método de cálculo para distâncias inferiores à referência, sendo a sua grande diferença causada pelo facto de não ser conhecido o valor da base total do triângulo rectângulo que intercepta o ponto de incidência do laser no objecto,  $k'$ , o ponto do emissor de laser e o ponto da câmara, como acontece no caso anterior. Tendo isto em consideração, o sistema de equações toma a seguinte forma:

$$\begin{cases} \tan(\Theta_L) = \frac{C'}{B} = \frac{C'}{B+B_{cam}} \\ \cos(90^\circ - \Theta_{Cam}) = \frac{B_{cam}}{D_h} \\ \tan(90^\circ - \Theta_{Cam}) = \frac{C'}{B_{cam}} \end{cases} \Leftrightarrow \begin{cases} C' = \tan(\Theta_L) \times (B + B_{cam}) \\ D_h = \frac{B_{cam}}{\cos(90^\circ - \Theta_{Cam})} \\ \tan(90^\circ - \Theta_{Cam}) = \frac{\tan(\Theta_L) \times (B + B_{cam})}{B_{cam}} \end{cases} \quad (4.6)$$

Observando o sistema de equações 4.6 e comparando-o com o previamente obtido no sistema 4.2 verifica-se, tal como era esperado, que a grande diferença se encontra na posição da base do triângulo  $B_{cam}$ , fazendo com que haja uma troca de sinais no relacionamento entre a base do laser e a base da câmara. Utilizando o mesmo método de cálculo da situação anterior e aplicando-o à situação presente, obtemos o sistema de equação final 4.7 que permite o cálculo directo das variáveis desconhecidas e da distância entre o ponto de incidência do laser no objecto e a câmara do sensor.

$$\begin{cases} C' = B \tan(\Theta_L) + \frac{B \tan(\Theta_L)^2}{\tan(90^\circ - \Theta_{Cam}) - \tan(\Theta_L)} \\ B_{cam} = \frac{B \tan(\Theta_L)}{\tan(90^\circ - \Theta_{Cam}) - \tan(\Theta_L)} \\ D_h = \sqrt{(C')^2 + (B_{cam})^2} \end{cases} \quad (4.7)$$

#### 4.2.2 Cálculo do Ângulo da Câmara

De forma a se poderem realizar os cálculos das distâncias referidos na secção 4.2.1, é necessário, a partir das imagens obtidas, saber qual a relação existente entre os pixels da imagem e o ângulo criado entre esse pixel e o eixo central da lente da câmara. Sendo esses ângulos não variantes com a distância, a sua evolução deveria ser linear. De forma a se tentar aproximar a evolução angular da imagem, foi criada uma maquete com uma representação da evolução angular com espaçamento em intervalos de  $10^\circ$ , para que seja possível retirar amostras das imagens capturadas pela câmara e aproximar a sua evolução angular. Na figura 4.7 é apresentada a maquete utilizada para o estudo da evolução angular da lente da câmara do sensor.

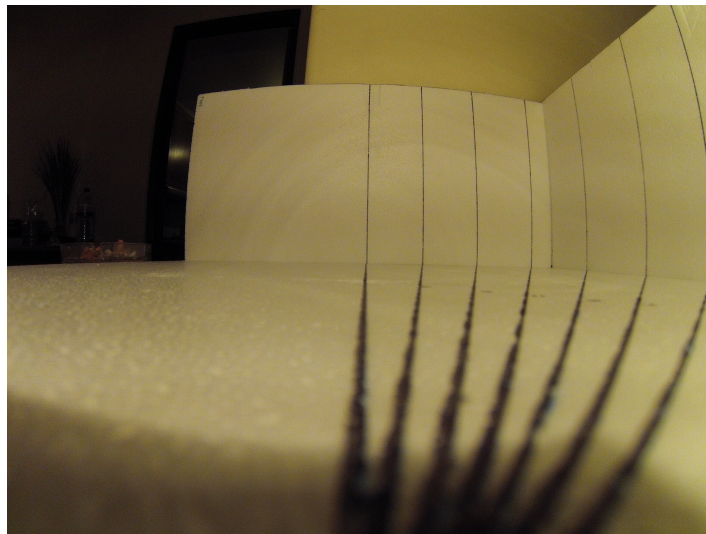


Figura 4.7: Maquete utilizada para o estudo da evolução angular da lente da câmara.

#### 4. VALIDAÇÃO

Como é possível verificar pela imagem, a lente da câmara utilizada deforma os objectos, sendo visível que, quanto mais afastado do centro da imagem, maior a sua deformação. Devido à existência desta deformação de imagem causada pela lente da câmara, a evolução angular já não se apresentará como linear. Retirando como amostra algumas linhas da figura 4.7 e utilizando a ferramenta de aproximação polinomial do *matlab*, *sftool*, dedicada a aproximar superfícies com base em amostras, foi possível obter uma aproximação polinomial da evolução angular da lente da imagem. Essa aproximação está representada na figura 4.8.

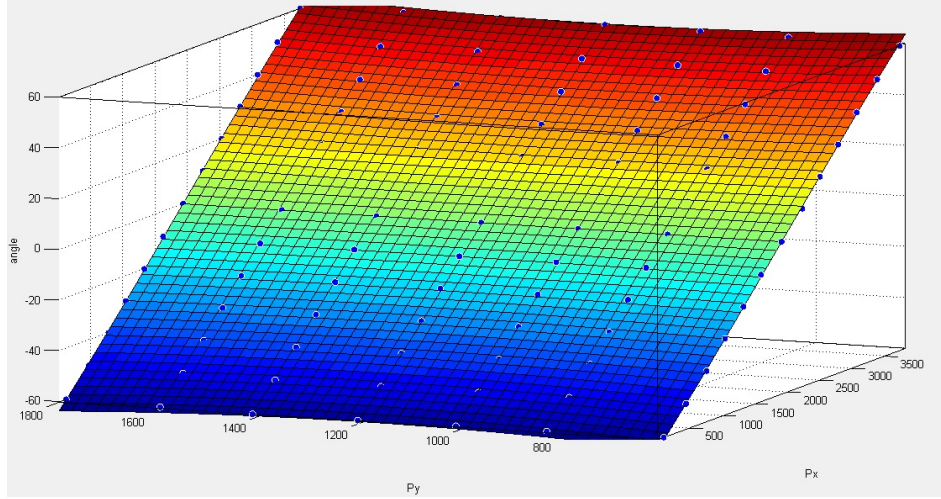


Figura 4.8: Superfície que aproxima a evolução angular da lente da câmara em relação às coordenadas dos pixels da imagem.

A equação da superfície apresentada na figura 4.8 pode ser definida pela equação 4.8:

$$\begin{aligned} p00 &= -75,70000 \\ p10 &= +00,03729 \\ p01 &= +00,02481 \\ p20 &= +01,66800 \times 10^{-06} \\ p11 &= -01,29200 \times 10^{-05} \\ p02 &= -01,01000 \times 10^{-05} \\ p30 &= -02,89500 \times 10^{-10} \\ p21 &= +07,62200 \times 10^{-24} \\ p12 &= +05,26200 \times 10^{-09} \\ p03 &= -08,65200 \times 10^{-23} \end{aligned} \tag{4.8}$$

$$\begin{aligned} angle(x, y) &= p00 + p10.x + p01.y + p20.x^2 + p11.x.y + p02.y^2 + p30.x^3 + \\ &\quad + p21.x^2.y + p12.x.y^2 + p03.y^3 \end{aligned}$$

A equação 4.8 apenas define a evolução dos ângulos da lente da câmara de uma forma horizontal, para se realizar de uma forma vertical, repete-se o estudo mas com a câmara posicionada sob a forma vertical na maquete. Obteve-se assim a equação 4.9, equação polinomial da superfície que aproxima a evolução dos ângulos verticais da câmara.

$$\begin{aligned}
p00 &= +76,47000 \\
p10 &= -00,03466 \\
p01 &= -00,05359 \\
p20 &= +01,00500 \times 10^{-05} \\
p11 &= +02,40700 \times 10^{-05} \\
p02 &= +05,08600 \times 10^{-07} \\
p30 &= +02,66700 \times 10^{-23} \\
p21 &= -06,98100 \times 10^{-09} \\
p12 &= +05,72400 \times 10^{-24} \\
p03 &= -01,17700 \times 10^{-10}
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
angle(x, y) &= p00 + p10.x + p01.y + p20.x^2 + p11.x.y + p02.y^2 + p30.x^3 + \\
&\quad + p21.x^2.y + p12.x.y^2 + p03.y^3;
\end{aligned}$$

As equações 4.8 e 4.9 permitem, através das coordenadas dos pixels  $Px$  e  $Py$  determinar respectivamente o ângulo horizontal e vertical que o ponto faz com a câmara.

### 4.3 Experimentação

De modo a ser possível a realização dos testes de validação do cenário proposto em 4.1, foi necessário todo um processo de experimentação de forma a testar cada um dos processos utilizados pelo protótipo. Estes testes, que foram sendo realizados ao longo do desenvolvimento do protótipo, tiveram como principal objectivo a determinação das suas limitações de forma a identificar possíveis pontos de melhoria.

A experimentação foi dividida em três testes distintos com objectivos diferentes. Iniciou-se pela realização de um teste de calibração para determinar e otimizar dados necessários ao cálculo da profundidade do cenário, em seguida foi realizado o teste de percepção de profundidade de uma linha laser, que consiste no teste de funcionamento do princípio básico de aquisição do protótipo e na criação de uma *point-cloud* necessária para a determinação "multiponto" de profundidades do cenário ao longo de uma linha vertical, por fim, foi realizado um teste de percepção tridimensional por rotação do sensor,

#### 4. VALIDAÇÃO

permitindo testar a viabilidade da utilização de múltiplas aquisições como as utilizadas no segundo teste, de forma a construir uma *point cloud* mais complexa e capaz de perceber não apenas uma linha vertical do cenário mas várias linhas verticais de forma a estudar todo o volume em redor do sensor.

##### 4.3.1 Teste de Calibração

Como já foi referido anteriormente, para se proceder ao cálculo das distâncias, é necessário encontrar valores constantes das equações de cálculo através da calibração. Nesta experiência foram executados vários testes de forma a determinar o valor do ângulo que o emissor de laser faz com a câmara. Para isso foram capturadas 7 imagens distintas e medidas as distâncias entre a câmara do sensor e o ponto central do laser emitido a uma altura constante. Na tabela 4.1 está apresentado o resultado final da experiência para um ângulo do laser emitido,  $\theta_L$ , de  $84,04^\circ$ .

Tabela 4.1: Resultados do Teste de Calibração

Medição nº	Px	Py	$\theta_{CAM}$ ( $^\circ$ )	Dist. Real (m)	Dist. Calc.(m)	Erro Relativo (%)
1	1951	1440	1,0249	1,165	1,1561	0,764
2	1950	1440	0,9921	1,155	1,1485	0,563
3	2058	1440	4,5338	3,980	3,9961	0,405
4	2051	1440	4,3043	3,351	3,4423	2,725
5	1930	1440	0,3362	1,016	1,0149	0,108
6	2005	1440	2,796	1,784	1,802	1,009
7	1915	1440	-0,1558	0,925	0,9336	0,930

Em cada medição da experiência foi seleccionado manualmente o pixel atravessado pela linha laser no centro vertical da imagem. Ao introduzir as coordenadas (Px,Py) desse pixel como entrada do polinómio de aproximação da evolução angular da lente da câmara, foi possível obter o ângulo que esse ponto faz com a câmara,  $\theta_{CAM}$ . Esse resultado serve como input do sistema de cálculo de distâncias que, usando as metodologias anteriormente referidas, calcula a distância entre a câmara e o ponto. O valor de  $\theta_L$  escolhido foi  $84,04^\circ$ , pois, verificando o erro relativo resultante de cada uma das medições, apresentou os valores mais baixos, com uma média inferior a 1%.



### 4.3.2 Teste de percepção de profundidade de uma linha laser

Nesta fase de experimentação, foram realizados vários testes com o objectivo de extrair das imagens captadas pela câmara a profundidade e altura de cada dos um pixels atravessados pela linha laser em relação à posição da câmara aquando da realização da aquisição. O objectivo principal desta experiência é o teste dos algoritmos de triângulação referidos na secção 4.2 e a sua capacidade de criação de uma *point-cloud* que se aproxime da realidade dos sistemas testados.

Seguindo o processo de aquisição de profundidade já referido anteriormente no capítulo 3, primeiro é realizada a aquisição da imagem que é depois transferida para o computador, para que possa ser processada. A imagem obtida pela câmara do sensor está representada na figura 4.9.

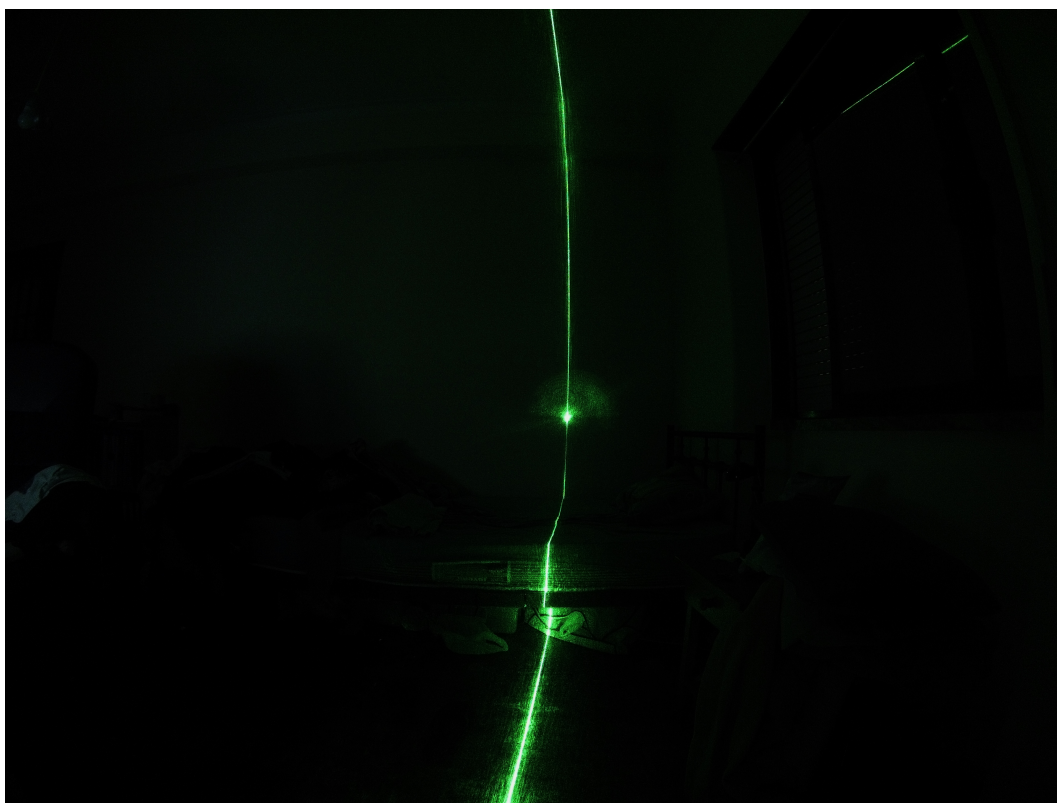


Figura 4.9: Imagem sem tratamento obtida no teste de percepção de profundidade de uma linha laser.

Obtida a imagem por parte do computador é iniciada agora a fase de processamento de imagem, em que se extrai todo o conteúdo perceptual necessário para o cálculo das coordenadas dos pontos da *point-cloud*. O output desta fase é um *array* com dimensão igual à altura da imagem, 2880 posições. Cada uma das posições do *array* contem a localização na qual o laser está a atravessar a linha correspondente de pixels.

#### 4. VALIDAÇÃO

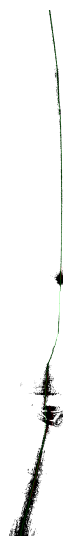


Figura 4.10: Resultado do processamento de selecção da linha laser.

Na figura 4.10 está representado o resultado do processamento de selecção da linha laser apresentada na figura 4.9 original. A fase final desta experiência vem em seguimento da aquisição dos pontos do *array* acima referidos. Baseando-se no cálculo das coordenadas cartesianas  $(X,Y,Z)$ , tal como calculado na secção 4.2, pretende-se agora criar uma representação gráfica da profundidade a que os objectos incididos pelo laser se apresentam em relação ao eixo de origem traçado pelo centro geométrico do sensor. Na figura 4.11 está apresentado o resultado desse cálculo, sendo possível observar a capacidade do protótipo de calcular uma aproximação da profundidade de uma linha laser sob a forma de uma *point-cloud*.

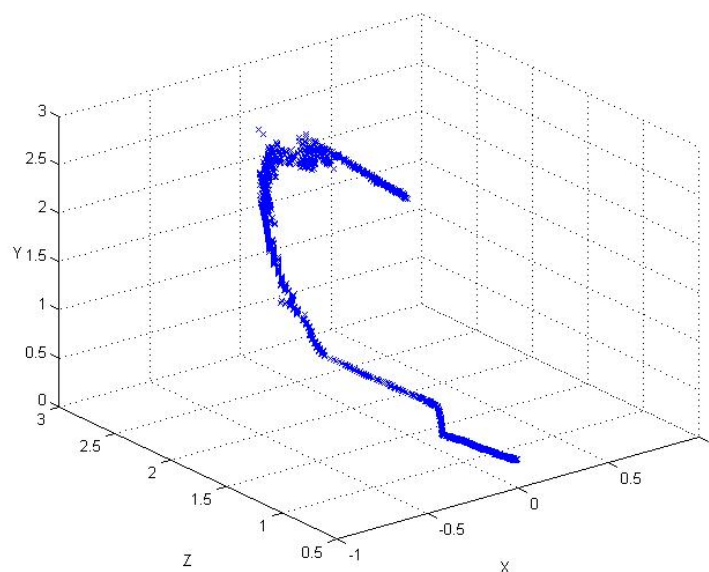
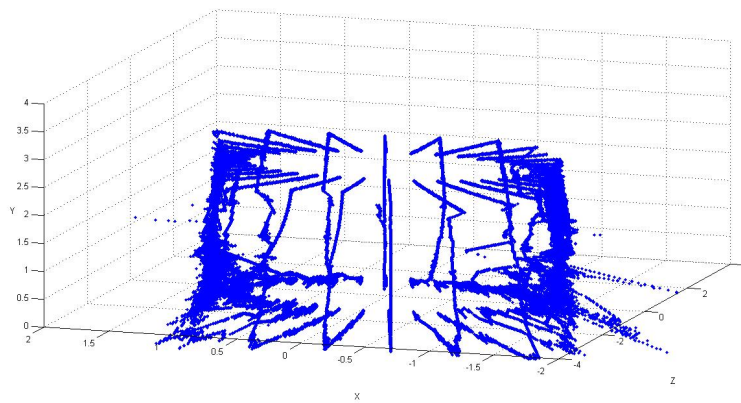


Figura 4.11: Representação tridimensional de uma linha laser.

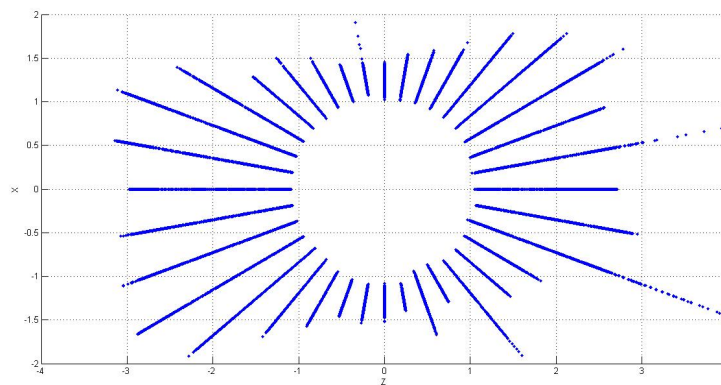
### 4.3.3 Teste de percepção tridimensional por rotação do sensor

Nesta fase de experimentação, foi realizado um teste com o objectivo de concluir sobre a capacidade do sensor de obter os dados necessários à percepção tridimensional de um espaço fechado, nomeadamente de uma garagem. Este teste consiste na validação da possibilidade de juntar várias aquisições singulares, tal como realizado no teste anterior, de forma a criar uma *point-cloud* que aproxime todo o espaço em redor do sensor.

Neste teste, foram realizadas 36 aquisições singulares espaçadas com uma amplitude de rotação de  $10^\circ$  no sentido horário do sensor sobre o eixo vertical dos YY. Nas figuras 4.12a e 4.12b é possível verificar os resultados da *point-cloud* obtida como resultado do teste realizado, sendo que a primeira é uma representação em perspectiva da garagem, vista de um ponto exterior e a segunda, uma vista da planta da garagem.



(a) Vista de Perspectiva.



(b) Vista de Planta.

Figura 4.12: *Point-cloud* resultante da aquisição de um espaço por rotação do sensor representadas em Perspectica (a) e Planta (b).

#### 4. VALIDAÇÃO

É possível verificar nas imagens que a *point-cloud* obtida aproxima uma estrutura cúbica rectangular tal como era pretendido provar com este teste.

### 4.4 Sinopse

Neste capítulo, foram tratados os métodos utilizados para a validação do protótipo, um estudo profundo sobre os componentes do protótipo criado e o seu funcionamento, e por fim, a realização de testes práticos de forma a não só se proceder à validação do processo proposto no capítulo anterior, como também testar as capacidades perceptuais do protótipo criado.

# CAPÍTULO 5

## Conclusões e Trabalhos Futuros

---

5.1	Conclusão . . . . .	55
5.2	Trabalhos Futuros . . . . .	56

---

### 5.1 Conclusão

Ao longo deste trabalho foi possível verificar que ainda não existe o método ideal para aquisição e percepção de dados tridimensionais. Existem contudo métodos que apresentam melhor performance em determinados cenários. A aquisição e percepção tridimensional são áreas ainda em crescimento e por essa razão, é necessário que haja uma inovação no estudo de novas metodologias e a continuidade do desenvolvimento das já existentes.

Apesar de se ter conseguido com êxito cumprir os objectivos propostos nesta dissertação, foi possível verificar a existência de alguns pormenores, no desenvolvimento do sistema de percepção espacial, cuja revisão poderiam melhorar substancialmente os resultados obtidos na validação desta dissertação. Alguns desses pormenores são nomeadamente: a utilização de uma linha laser mais fina e com menos ruído luminoso, a utilização de um motor de passo que realizasse a rotação em intervalos angulares menores, de forma a criar uma *cloud-point* mais densa horizontalmente.

A calibração do sensor revelou-se ser uma das tarefas mais importante e difícil de todo o processo, já que foi necessário criar métodos para obter valores com o menor erro possível: a distância existente entre a câmara e o emissor laser, ângulo criado entre emissor laser e a base do sensor e, por fim, o estudo da deformação angular da lente. Cada um destes factores é responsável pela introdução de erros de medição na percepção espacial que agravam os erros de medição ao se distanciarem dos seus valores reais.

## 5.2 Trabalhos Futuros

Tendo em mente os aspectos referidos na *Conclusão*, os trabalhos futuros propostos baseiam-se principalmente no ramo de optimização do processo de calibração, nomeadamente a utilização de lentes de *fresnel* como alternativa ao prisma utilizado para produção de uma linha laser e melhoria dos processos de calibração.

Outra possibilidade seria a aplicação deste sensor a *drones* com o objectivo de obter automaticamente a percepção, sendo necessário estudar e desenvolver a capacidade algorítmica do aparelho, de forma a que este forneça pontos de referência ao sensor, para que este se possa situar correctamente no espaço e colocar os pontos da *point-cloud* no sítio correcto.

Um outro projecto que melhoraria significativamente a qualidade dos dados de percepção do protótipo construído seria a criação de um algoritmo de avaliação da *point-cloud* criada pelo sensor, de forma a detectar erros e proceder a correcções da mesma, realizando no final uma interpolação entre os vários pontos da *point-cloud* de forma a criar representações dos objectos.

# Bibliografia

- [1] D. Barber and J. Mills. *3D Laser Scanning for Heritage (second edition)*. English Heritage, 2011.
- [2] N. E. Beckman, K. Bierhoff, and J. Aldrich. Verifying correct usage of atomic blocks and typestate. *SIGPLAN Not.*, 43(10):227–244, 2008.
- [3] P. Chmela and M. Dobrovoln. The optical measuring device for the autonomous exploration and mapping of unknown environments. VII(4):41–50, December 2012.
- [4] Coord3 Industries Srl. Coordinate measuring machine history – fifty years of cmm history leading up to a measuring revolution. <http://www.coord3-cmm.com/50-years-of-coordinate-measuring-machine-industry-developments-and-history/>. Accessed: 20/08/2014.
- [5] B. Curless. From range scans to 3d models. *ACM SIGGRAPH Computer Graphics*, 33(4):38–41, November 1999.
- [6] E. R. Davies. *Computer and Machine Vision: Theory, Algorithms and Practicalities*. Elsevier Inc., 2012.
- [7] GitHub. A lightweight python class for interfacing with multiple gopros. <https://github.com/joshvillbrandt/GoProController>. Accessed: 09/08/2014.
- [8] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (second edition)*. Prentice Hall, 2002.
- [9] Google. Acerca do street view. <http://www.google.com/intl/pt/maps/about/behind-the-scenes/streetview/>. Accessed: 19/09/2014.
- [10] GoPro. Gopro hero3 silver edition technical specs. [pt.gopro.com//cameras/hd-hero3-silver-edition#technical-specs](http://pt.gopro.com//cameras/hd-hero3-silver-edition#technical-specs). Accessed: 10/08/2014.
- [11] Hack a Day. Gopro hero 3 black edition python api/wrapper. <http://forums.hackaday.com/viewtopic.php?f=3&t=3336>. Accessed: 09/08/2014.
- [12] Hack a Day. Gopro wifi bacpac control. <http://www.tavarn.org/GoProWEBApp/GoPro-WiFi-bacpac-control.html>. Accessed: 09/08/2014.
- [13] K. C. J, C, A. K, D, L. D, D, and T. W, F. Performance analysis of low-cost triangulation-based camera: Microsoft kinect system. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX(B5):175–180, 2012.
- [14] K. Khoshelham and S. O. Elberink. Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors (Basel, Switzerland)*, 12(2):1437–1454, January 2012.

- [15] L. Li. Time-of-flight camera – an introduction. *Texas Instruments*, January 2012.
- [16] C. Mythili. Efficient technique for color image noise reduction. *The Research Bulletin of Jordan ACM*, II(iii):41–44, 2011.
- [17] M. D. Sutton, I. A. Rahman, and T. . Garwood, Russell J’.